

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

CARACTÉRISATION DE LA TRANSCRIPTION ANTISENS CHEZ LES RÉTROVIRUS
HUMAINS HTLV-2, HTLV-3 ET HTLV-4

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN BIOLOGIE

PAR
MARILÈNE HALIN

MARS 2009

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens d'abord à remercier mon directeur de recherche pour son accueil, sa patience et ses précieux conseils. Lors de ces deux années, il a su m'encadrer tout en me laissant suffisamment de liberté, me permettant ainsi de cheminer dans un environnement de confiance et d'apprentissage. De plus, le projet qui m'a été attribué a su entretenir mon intérêt, me permettant ainsi de passer deux superbes années.

Je souhaite également remercier tous les membres du laboratoire du Dr. Benoît Barbeau pour leurs conseils et leur enthousiasme, qui ont assurément rendu ce séjour inoubliable. Un merci particulier à Sébastien Landry qui a passé beaucoup de temps à m'expliquer les rudiments de la biologie moléculaire et sans qui je ne n'aurais certainement pas l'esprit aussi critique. Et merci à vous groupe : Chirine Toufaily, Émilie Larocque et Amandine Vargas.

Je voudrais également saluer ma famille et mon copain et leur dire que c'est grâce à eux si je me lève tous les matins le cœur léger et le sourire aux lèvres.

Finalement, merci au département des sciences biologiques de l'Université du Québec À Montréal et au centre de recherche BioMed.

TABLE DES MATIÈRES

LISTE DES FIGURES.....	vii
LISTE DES TABLEAUX.....	ix
LISTE DES ABRÉVIATIONS.....	x
LISTE DES SYMBOLES.....	xiv
RÉSUMÉ.....	xv
INTRODUCTION.....	1
CHAPITRE I.....	3
ÉTAT DES CONNAISSANCES.....	3
1.1 Les rétrovirus.....	3
1.1.1 Organisation génomique.....	4
1.1.2 Cycle de réplication.....	4
1.2 Les rétrovirus HTLV.....	7
1.2.1 Origine et distribution des différents membres.....	7
1.2.2 Pathogénicité et transmission.....	8
1.2.3 Structure.....	10
1.2.4 Mécanisme d'infection.....	11
1.2.5 Organisation génomique.....	12
1.2.5.1 Tax et Rex.....	14
1.3 La transcription.....	16
1.3.1 Polyadénylation.....	18
1.3.2 Épissage.....	18
1.4 Transcription antisens.....	20
1.4.1 Chez les organismes procaryotes et eucaryotes.....	20
1.4.2 Chez les rétrovirus.....	21
1.4.3 Chez HTLV-1.....	21

1.4.3.1	Structure du transcrit.....	23
1.4.3.2	« HTLV-1 bZIP factor »	23
1.4.3.2.1	Structure moléculaire	24
1.4.3.2.2	Activité promotrice.....	25
1.4.3.2.3	Localisation cellulaire	26
1.4.3.2.4	Fonctions	27
1.4.3.2.5	HBZ et ATL	31
1.5	Hypothèse de travail	31
CHAPITRE II.....		33
ARTICLE 1		33
2.1	Contribution des auteurs	33
2.2	Résumé.....	34
2.3	Human T-cell Leukemia Virus Type 2 produces a spliced antisense transcript encoding a protein that lacks a bZIP domain but inhibits Tax2-mediated transcription....	35
2.3.2	Introduction	37
2.3.3	Material and methods	40
2.3.3.1	Cell lines	40
2.3.3.2	HTLV-2 carrier samples	40
2.3.3.3	Plasmids	40
2.3.3.4	Transfections.....	41
2.3.3.5	RT-PCR, 5' and 3' RACE analyses.....	41
2.3.3.6	RT-PCR analyses in infected cell lines.....	42
2.3.3.7	Northern blot analyses	42
2.3.3.8	Generation of APH2-antibodies.....	43
2.3.3.9	Western blot analyses	43
2.3.3.10	Immunofluorescence analyses.....	44
2.3.3.11	Luciferase assays.....	44
2.3.4	Results	46
2.3.4.1	Existence of a putative ORF in the HTLV-2 minus strand.....	46
2.3.4.2	Detection of a spliced APH-2 transcript in HTLV-2 transfected cells.	46

2.3.4.3	<i>APH-2</i> mRNA is present in HTLV-2 infected cells and in HTLV-2 carriers in vivo.....	48
2.3.4.4	HTLV-2 antisense transcription initiates at different positions in the 3'LTR.	48
2.3.4.5	Position of the poly A site.....	49
2.3.4.6	<i>APH-2</i> and <i>HBZ-SP1</i> do not have a similar localization in transfected cells.	50
2.3.4.7	<i>APH-2</i> down-regulates Tax2-mediated HTLV-2 LTR activity.	50
2.3.5	Discussion	52
2.3.6	Acknowledgments.....	55
2.3.7	Figure legends.....	56
2.3.8	Figures.....	60
CHAPITRE III		75
ARTICLE 2		75
3.1	Contribution des auteurs	75
3.2	Résumé.....	76
3.3	HTLV-3 and HTLV-4 viruses produce antisense transcripts encoding proteins with a Tax-inhibiting function	77
3.3.1	Abstract	78
3.3.2	Introduction	79
3.3.3	Material and methods.....	81
3.3.3.1	Cell lines and antibodies	81
3.3.3.2	Plasmids	81
3.3.3.3	Transfection	83
3.3.3.4	RT-PCR and 3'RACE analyses	84
3.3.3.5	Confocal microscopy	85
3.3.3.6	Western blot analysis.....	85
3.3.4	Results.....	86
3.3.4.1	Detection of spliced antisense transcripts in HTLV-3 and HTLV-4-transfected cells.....	86

3.3.4.2	<i>APH-3</i> and <i>APH-4</i> transcripts are both polyadenylated.....	87
3.3.4.3	Amino acid sequence of <i>APH-3</i> and <i>APH-4</i> and their cellular localization. 87	
3.3.4.4	Promoter studies on both HTLV-3 and HTLV-4 antisense transcripts.....	89
3.3.4.5	Tax-driven HTLV LTR activation is inhibited by <i>APH-3</i> and <i>APH-4</i> expression.....	89
3.3.5	Discussion	91
3.3.6	Acknowledgments.....	94
3.3.7	Figure legends	95
CHAPITRE IV		112
DISCUSSION COMPLÉMENTAIRE AUX ARTICLES ET CONCLUSIONS		112
4.1	Les rétrovirus HTLV-2, -3 et -4 expriment un transcrit antisens.....	112
4.2	L'initiation de la transcription a lieu dans le LTR3' et les transcrits sont polyadénylés.	115
4.3	L'épissage des transcrits engendre des cadres de lecture ouverts distribués sur deux exons.	117
4.4	<i>APH-2</i> , -3 et -4 se localisent dans le noyau en plus d'être présents dans le cytoplasme pour <i>APH-3</i>	117
4.5	<i>APH-2</i> , -3 et -4 inhibent la transcription de HTLV-1 dépendante de la protéine virale Tax.	120
4.6	Le motif leucine zipper ne semble pas présent chez <i>APH-2</i> , -3 et -4.	120
4.7	La transcription antisens est modulable par des agents activateurs des cellules T. 122	
4.8	Conclusion	123
RÉFÉRENCES.....		125

LISTE DES FIGURES

Figure 1.1 : Représentation des étapes d'infection, d'intégration et de réplication d'un rétrovirus dans une cellule.....	5
Figure 1.2 : Résultat de la transcription inverse de l'ARN simple brin en ADN double brins.....	6
Figure 1.3 : Cellule sanguine d'un patient atteint d'ATL présentant un lymphocyte T à noyau multilobulé.....	9
Figure 1.4 : Structure du rétrovirus HTLV-1.....	11
Figure 1.5 : Génome et structures des transcrits des rétrovirus HTLV-1 et HTLV-2.....	13
Figure 1.6 : Représentation du complexe d'initiation de la transcription chez les organismes eucaryotes.....	17
Figure 1.7 : Schéma des étapes d'épissage.....	19
Figure 1.8 : Représentation des séquences nucléotidiques et protéiques des différents isoformes de HBZ.....	22
Figure 1.9 : Structure moléculaire de la protéine HBZ.....	24
Figure 1.10 : Structure tridimensionnelle d'un <i>leucine zipper</i> interagissant avec l'ADN.....	25
Figure 1.11 : Localisation cellulaire des différents isoformes de HBZ dans les cellules COS-7 par microscopie confocale.....	27
Figure 1.12 : Schéma des différentes actions de HBZ et conséquences sur la transcription.....	30
Figure 2.1 : HTLV-2 antisense transcripts are detected and are spliced.....	60-62
Figure 2.2 : The splice donor and acceptor sequences of the APH-2 transcript are conserved.....	63-64
Figure 2.3 : The spliced APH-2 mRNA is detected in HTLV-2-infected cell lines and in PBMCs from infected blood donors.....	65
Figure 2.4 : HTLV-2 antisense transcription initiates in the 3' LTR.....	66
Figure 2.5 : Identification of the poly A addition site of the APH-2 transcript.....	67-68

Figure 2.6 : Amino acid sequence of the APH-2 protein.....	69
Figure 2.7 : APH-2 localizes to the nucleus.....	70-71
Figure 2.8 : APH-2 represses Tax2 dependent transcription from the HTLV-2 LTR.	72-73
Figure 2.9: Comparison of HBZ and APH-2 transcripts.....	74
Figure 3.1 : Position of the antisense ORF in the HTLV-3 and HTLV-4 genomes.....	98-99
Figure 3.2 : HTLV-3 and HTLV-4 antisense transcripts are detected and are spliced...100-102	
Figure 3.3 : Identification of the poly A addition site for both <i>APH-3</i> and <i>APH-4</i> transcripts.....	103-104
Figure 3.4 : Amino acid sequence of the APH-3 and APH-4 proteins.....	105
Figure 3.5 : APH-3 and APH-4 demonstrate different cellular localizations.....	106-108
Figure 3.6 : Antisense expression is upregulated in Jurkat T cells after activation.....	109-110
Figure 3.7 : APH-3 and APH-4 repress Tax1 and Tax3 dependent transcription from the HTLV-1 LTR.....	111

LISTE DES TABLEAUX

Tableau 1.1 : Classification des rétrovirus selon le comité international de taxonomie des virus.....	4
---	---

LISTE DES ABRÉVIATIONS

A : adénine

AD : *activation domain*

ADN : acide désoxyribonucléique

ADNc : acide désoxyribonucléique complémentaire

AP-1 : *activator protein 1*

APH : *antisense protein of HTLV*

ARN : acide ribonucléique

ARNm : acide ribonucléique messenger

ATF : *activating transcription factor*

ATL : *adult T-cell leukemia*

ATLL : *adult T-cell leukemia/lymphoma*

ATP: adenosine-5'-triphosphate

BBP : *branch point binding protein*

BER : *base excision repair*

bp: *base pair*

bpV[pic]: *dipotassium bisperoxo (picolinate) oxovanadate*

BR1/BR2: *basic regions*

bZIP : *basic leucine zipper*

C: cytosine

CBP/p300 : *CREB binding protein / E1A binding protein p300*

CD4 / 8: *cluster of differentiation 4 / 8*

cDNA: *complementary deoxyribonucleic acid*

CPSF: *cleavage and polyadenylation specificity factor*

CRE : *cAMP response element*

CREB: *cAMP response element-binding*

CstF: *cleavage stimulatory factor*

DBD: *DNA-binding domain*

DMEM: *Dulbecco's Modified Eagle's Medium*

dNTP: désoxyribonucléotide triphosphate

DPE: *downstream promoter element*
 DSBR: *DNA double-strand break repair*
 DTT: dithiothreitol
 ECL: *electrochemiluminescence*
 EDTA: acide éthylènediaminetétraacétique
 env: envelope
 ESE: *exonic splicing enhancer*
 ESS: *exonic splicing silencer*
 G: guanine
 gag : *group-specific antigen*
 GFP: *green fluorescent protein*
 GTP: *guanosine-5'-triphosphate*
 HAM/TSP: *HTLV-1-associated myelopathy /tropical spastic paraparesis*
 HBZ: *HTLV-1 bZIP factor*
 HSPG: *heparan sulfate proteoglycans*
 hTERT: *human telomerase reverse transcriptase*
 HTLV: *human T-cell lymphotropic virus*
 IL: interleukine
 KLH: *keyhole limpet hemocyanin*
 LTR: *long terminal repeat*
 $(\text{MgCO}_3)_4 \cdot \text{Mg}(\text{OH})_2 \cdot 5\text{H}_2\text{O}$: magnésium hydroxycarbonate 5-hydrate
 MgSO_4 : sulfate de magnésium
 MMR: *mismatch repair*
 MOPS: acide 3-(N-morpholino)propanesulfonique
 mRNA: *messenger ribonucleic acid*
 NaCl: chlorure de sodium
 NaF: fluorure de sodium
 Na_2PO_4 : tétraoxophosphate de disodium
 NER: *nucleotide excision repair*
 $\text{NF}\kappa\text{B}$: *nuclear factor-kappa B*
 NLS: *nuclear localisation signal*

NP-1: *neuropilin-1*

NP-40: *nonyl phenoxy polyethoxylethanol*

nt: nucléotide

ORF: *open reading frame*

PABP: *poly-A-binding proteins*

PAP: *poly A polymerase*

pb: paire de bases

PBMC: *peripheral blood mononuclear cell*

PBS: *phosphate buffered saline*

PCNA: *proliferating cell nuclear antigen*

PCR: *polymerase chain reaction*

PHA-P: *phytohemagglutinin*

PI: *propidium iodide*

PIC: *pre-initiation complex*

PMA: *phorbol myristate acetate*

PML: *promyelocytic leukemia*

pol: polymérase

poly A: polyadénylation

PTP: *protein tyrosine phosphatase*

pro: protéase

PVDF: *polyvinylidene fluoride*

PTLV: *primate T-cell lymphotropic virus*

q-PCR: *quantitative polymerase chain reaction*

q-RT-PCR: *quantitative reverse transcription polymerase chain reaction*

R: région unique répétée

RACE: *rapid amplification of cDNA ends*

RexRE: *Rex response element*

RLU : *relative light unit*

RNA: *ribonucleic acid*

RPMI : *roswell park memorial institute*

RT: *rétro-transcriptase*

SA : site accepteur d'épissage
 SD : site donneur d'épissage
 SDS : dodécyl sulfate de sodium
 SDS-PAGE : *sodium dodecyl sulfate polyacrylamide gel electrophoresis*
 snRNP : *small nuclear ribonucleoproteins*
 SRF : *serum response factor*
 STLV : *simian T-cell lymphotropic virus*
 SU : glycoprotéine de surface
 SUMO : *small ubiquitin-like modifier*
 T : thymine
 TBP : *TATA binding protein*
 TCR : *T cell receptor*
 TFII : *transcription factor of RNA polymérase II*
 TM : glycoprotéine transmembranaire
 TNF α : *tumor necrosis factor alpha*
 TRE : *Tax responsive elements*
 Tris-HCl : hydrochlorure de 2-amino-2-(hydroxyméthyl)-1,3-propanediol
 U : uracile
 U2AF : *U2 small nuclear ribonucleoprotein particle auxiliary factor*
 U3 / U5 : séquences uniques répétées
 usHBZ : *unspliced HBZ*
 UV : ultraviolet
 VIH-1: virus de l'immunodéficience humaine

LISTE DES SYMBOLES

AA : acide aminé

h : heure

kb : kilobase

kDa : kilodalton

mg : milligramme

µg : microgramme

µM : micromolaire

min : minute

ml : millilitre

mM : millimolaire

ng : nanogramme

nm : nanomètre

U : unité

V : volt

α : alpha

β : bêta

% : pourcentage

≈ : environ

°C : degré Celsius

RÉSUMÉ

Chez les rétrovirus, la production des protéines virales dépend de l'expression d'un transcrit sens pleine longueur, initié dans le LTR5' et épissé de façon alternative. Des études ont récemment mis en évidence l'existence d'un nouveau type de transcrit antisens, initié dans le LTR3' de certains rétrovirus humains (VIH-1 et HTLV-1). Chez le virus HTLV-1, plusieurs études ont permis l'identification de deux formes d'épissage du transcrit antisens ainsi que la caractérisation d'une nouvelle protéine virale, nommée HBZ (*HTLV-1 bZIP factor*). Cette protéine se localise au noyau et régule négativement la transcription virale en interagissant avec des facteurs de transcription contenant un motif *leucine zipper*. Cette étude a pour objectif de caractériser la transcription antisens des rétrovirus humains HTLV-2, HTLV-3 et HTLV-4. Des études de RT-PCR ont permis de détecter un transcrit antisens chez des cellules 293T transfectées par l'ADN proviral des rétrovirus HTLV-2, -3 et -4. Le séquençage des signaux obtenus a confirmé la présence d'épissage dans chacun des transcrits étudiés. Des sites d'initiation de la transcription ont pu être identifiés par des analyses de 5' RACE sur un clone du virus HTLV-2. Un site fonctionnel de polyadénylation a été identifié chez un clone de ces trois rétrovirus. La comparaison des séquences de la protéine HBZ et des ORF antisens des rétrovirus HTLV-2, -3 et -4 démontre que le motif *leucine zipper* ne semble pas présent chez ces derniers. Dans le but de permettre l'expression et la détection des protéines antisens, les parties codantes des transcrits antisens des rétrovirus HTLV-2, -3 et -4 ont été clonées dans un vecteur d'expression contenant une étiquette Myc. L'analyse des cellules transfectées par microscopie confocale a permis de détecter la protéine antisens de HTLV-3 (APH-3) à la fois dans le noyau et dans le cytoplasme ainsi que les protéines antisens de HTLV-2 et -4 (APH-2 et APH-4) majoritairement dans le noyau. La co-transfection des vecteurs d'expressions de APH-2, APH-3 et APH-4 en présence des vecteurs exprimant les protéines virales Tax1, Tax2 ou Tax3 en plus du vecteur contenant le LTR de HTLV-1 ou de HTLV-2 en amont du gène de la luciférase dans des cellules 293T, a permis de mettre en évidence une inhibition de l'activation de la transcription virale dépendante de Tax par ces nouvelles protéines antisens. Finalement, afin de caractériser l'activité promotrice antisens de ces rétrovirus, le gène rapporteur de la luciférase a été inséré dans les deuxièmes exons de APH-3 et APH-4, compris dans les vecteurs contenant la portion 3' des génomes proviraux de HTLV-3 et -4. La transfection de ces constructions dans des cellules lymphocytaires humaines (Jurkat E6.1), suivie d'une stimulation par une série d'agents activateurs de ces dernières, a permis d'observer une modulation de la transcription antisens. Ces recherches portant sur la transcription antisens des rétrovirus humains ont permis de mettre en évidence que les transcrits antisens des rétrovirus HTLV-2, -3 et -4 ont la capacité de coder pour des nouvelles protéines virales. Les analyses de séquences effectuées ainsi que les observations en microscopie confocale laissent croire que ces protéines pourraient jouer des rôles différents de ceux précédemment attribués à la protéine HBZ. De plus amples études seront nécessaires afin d'identifier les rôles de ces protéines nouvellement découvertes dans le cycle de réplication viral.

MOTS CLÉS : APH, rétrovirus, HTLV, transcription antisens, HBZ.

INTRODUCTION

On a longtemps cru que, chez les rétrovirus humains, la production de la totalité des protéines virales dépendait de l'expression d'un transcrit sens unique, épissé de façon alternative. Ces dernières années, la mise en évidence de l'existence d'un transcrit antisens chez certains rétrovirus, tels que le VIH-1 et HTLV-1 (Cavanagh et al., 2006; Landry et al., 2007), a forcé les rétrovirologistes à reconsidérer la biologie moléculaire de ces virus. Récemment, de nombreuses études ont porté sur la transcription antisens du rétrovirus humain HTLV-1. Il est maintenant connu que ce dernier virus a la possibilité d'exprimer un transcrit antisens, lequel code pour une protéine nommée HBZ (*HTLV-1 bZIP factor*). Bien que le rôle de cette nouvelle protéine virale dans le cycle de réplication ne soit pas encore bien défini, plusieurs études récentes ont permis de démontrer l'impact d'HBZ sur l'activité d'autres protéines pouvant provenir à la fois du virus ou de la cellule hôte, telles que la protéine virale Tax ainsi que certains facteurs transcriptionnels cellulaires tels que CREB et AP-1. Les données acquises jusqu'à maintenant sur HBZ et son rôle dans la régulation de l'expression génique ne laissent d'autres choix que de considérer cette protéine comme un nouveau joueur important dans la réplication du virus HTLV-1. Notre groupe de recherche a présenté la caractérisation complète d'un transcrit antisens rétroviral, soit celui de HTLV-1. Il existe également trois autres rétrovirus de cette famille, soient HTLV-2, -3 et -4 et aucune recherche ne fait état d'une quelconque transcription antisens chez ces derniers. Les objectifs de cette étude sont la caractérisation de la transcription antisens chez ces rétrovirus en plus d'étudier l'activité promotrice, la localisation cellulaire et le potentiel inhibiteur de la transcription virale dépendante de Tax, des protéines codées par ces transcrits. D'abord, des expériences de RT-PCR et de RACE permettront la caractérisation des transcrits. Ensuite, des observations en microscopie confocale des protéines issues de cette transcription antisens seront effectuées. L'activité promotrice sera étudiée en faisant appel à des constructions contenant le gène de la luciférase inséré dans le cadre de lecture ouvert antisens des rétrovirus HTLV-3 et -4. Finalement des expériences de co-transfections en présence de ces nouvelles protéines, des protéines virales Tax et des vecteurs contenant le gène de la luciférase en aval des LTR rétroviraux permettront de vérifier si ces protéines antisens sont capables d'inhiber la transcription virale dépendante de Tax.

Dans un premier temps, une revue de la littérature sera présentée, suivie du contenu de deux articles scientifiques issus des résultats obtenus dans le cadre de ma maîtrise. Il s'en suivra une discussion complémentaire aux articles et une conclusion générale.

CHAPITRE I

ÉTAT DES CONNAISSANCES

1.1 Les rétrovirus

Autrefois, la recherche en biologie portait principalement sur l'analyse des symptômes occasionnés chez l'hôte plutôt que sur la compréhension des mécanismes cellulaires à l'origine des dérégulations lors d'une infection par un virus. Ce n'est que dans les années 1970 que la recherche a émergé au point de vue moléculaire, permettant ainsi la découverte et la caractérisation de plusieurs virus (Coffin et al., 1997).

La famille des rétrovirus réunit les virus à ARN enveloppés ayant une structure, une composition et un cycle de réplication semblables. Le diamètre des particules virales est généralement situé entre 80 et 100 nm. La structure des virions varie entre les différents membres, mais leur matériel génétique est assurément composé d'ARN simple brin (entre 7 et 12 kb), linéaire, non segmenté et de polarité positive. La principale caractéristique de cette famille de virus réside dans sa stratégie répllicative, qui implique une étape clé de transcription inverse de son génome d'ARN en ADN double brin, qui sera intégré dans le génome de la cellule hôte.

Dans son huitième rapport publié en juin 2005, l'organisme international de taxonomie des virus a sous-divisé la famille des *retroviridae* en deux sous-familles, soit les *orthoretrovirinae*, et les *spumaretrovirinae*, lesquelles contiennent sept genres (tableau 1.1).

Tableau 1.1 : Classification des rétrovirus selon le comité international de taxonomie des virus.

Famille	Sous-famille	Genre	Exemples d'espèces
<i>Retroviridae</i>	<i>Orthoretrovirinae</i>	<i>Alpharetrovirus</i>	<i>Avian leukosis virus</i>
		<i>Betaretrovirus</i>	<i>Mouse mammary tumor virus</i>
		<i>Gammaretrovirus</i>	<i>Murine leukemia virus</i>
		<i>Deltaretrovirus</i>	<i>Human T-cell lymphotropic virus I</i>
		<i>Epsilonretrovirus</i>	<i>Walleye dermal sarcoma virus</i>
		<i>Lentivirus</i>	<i>Human immunodeficiency virus I</i>
	<i>Spumaretrovirinae</i>	<i>Spumavirus</i>	<i>Simian foamy virus</i>

1.1.1 Organisation génomique

Il existe deux catégories de rétrovirus : simple et complexe. Ces premiers contiennent uniquement les gènes essentiels à leur réplication, soient les gènes *gag*, lequel code pour des protéines de la matrice, de la capsid et du nucléocapsid; *pro/pol*, qui engendrent la production d'une protéase, de la reverse transcriptase et de l'intégrase et *env*, qui produit les protéines de surface et transmembranaire de l'enveloppe. D'un autre côté, les rétrovirus complexes possèdent également le matériel génétique nécessaire à la production de protéines dites accessoires qui jouent des rôles dans la régulation de l'expression des gènes viraux et cellulaires. Notons que la production de ces dernières est le résultat d'épissages multiples.

1.1.2 Cycle de réplication

Comme mentionné précédemment, la principale particularité des rétrovirus réside dans leur cycle de réplication. En effet, ces derniers ont la capacité de s'intégrer dans le génome de la cellule hôte *via* une étape clé : la réaction de transcription inverse de leur matériel génétique, permettant ainsi de transformer leur génome constitué d'ARN simple brin en ADN double brin.

Le cycle de réplication des rétrovirus débute par l'attachement de la particule virale à des récepteurs membranaires spécifiques à la cellule cible (figure 1.1).

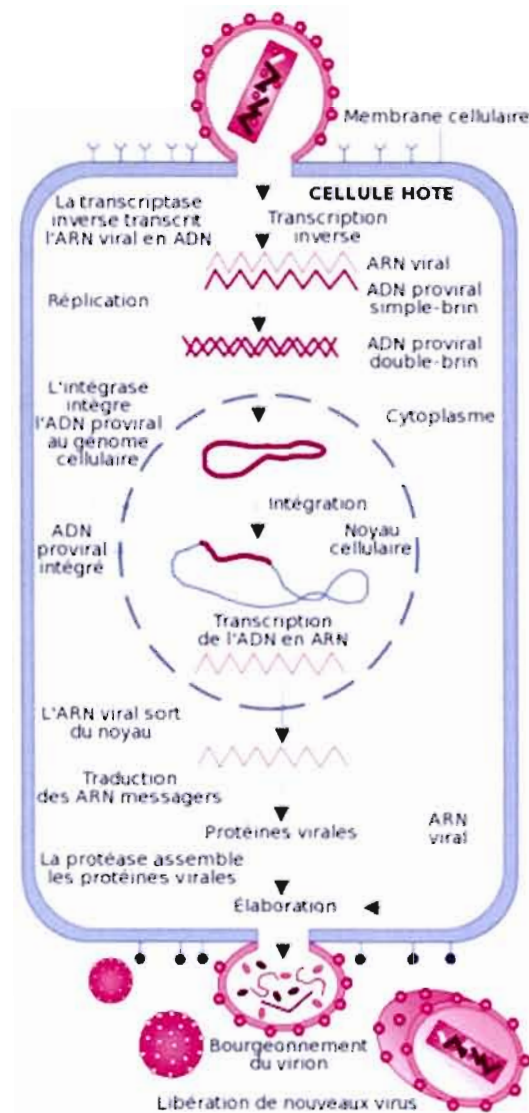


Figure 1.1 : Représentation des étapes d'infection, d'intégration et de réplication d'un rétrovirus dans une cellule. Les virions s'attachent à la membrane cellulaire et y pénètrent. L'ARN est rétrotranscrit en ADN qui est intégré dans l'ADN cellulaire et transcrit à nouveau pour produire les protéines virales et les virions. Image modifiée et tirée de : http://fr.wikipedia.org/wiki/Virus_de_l'immunod%C3%A9ficience_humaine.

La glycoprotéine présente à la surface du virus (SU) lie un récepteur spécifique sur la membrane de la cellule hôte, permettant ainsi la fusion des membranes avec l'aide de la seconde glycoprotéine (TM), située dans la membrane virale. Ces deux glycoprotéines virales qui sont à l'origine de la fusion membranaire, proviennent du clivage de la polyprotéine

codée par le gène de l'enveloppe en deux sous-unités distinctes. Le mécanisme exact par lequel ces dernières permettent la fusion des membranes reste à définir mais il nécessite assurément un changement de conformation.

Par la suite, l'ARN ainsi libéré dans le cytoplasme est rétrotranscrit en ADN simple brin, puis double brin, par l'enzyme transcriptase inverse présente dans les particules virales *via* une série de réactions de polymérase et de digestion par la ribonucléase H. Ces étapes de rétrotranscription engendrent la duplication des régions U5 et U3 de l'ARN génomique générant ainsi de l'ADN génomique encadré par deux LTR identiques (Figure 1.2).

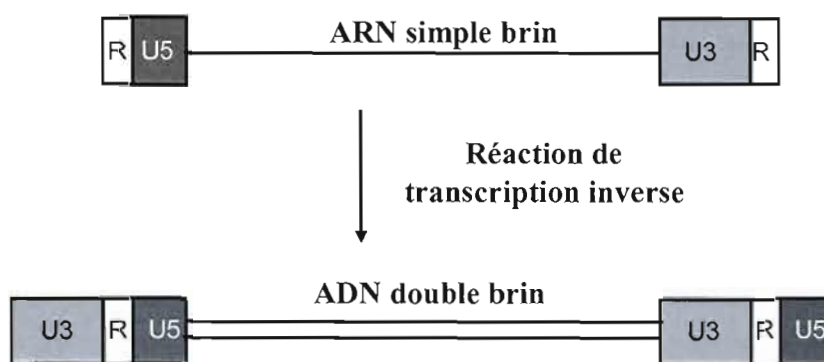


Figure 1.2 : Résultat de la transcription inverse de l'ARN simple brin en ADN double brin. Il y a duplication des régions U5 et U3 lors de la transcription inverse, ce qui est à l'origine des LTR, chacun constitué des régions U3, R et U5. Image modifiée de Coffin *et al.*, 1997.

Il s'ensuit le transport de l'ADN viral dans le noyau et son intégration dans le génome de la cellule hôte à l'aide de l'enzyme intégrase, elle aussi codée par le rétrovirus. Cette intégration dans le génome de l'hôte a lieu au hasard ou presque, puisqu'il a été démontré que certains rétrovirus ciblent préférentiellement certaines portions du génome telles les régions régulatrices et promotrices caractérisées par de fréquentes activités transcriptionnelles (Schoroder *et al.*, 2002 ; Mitchell *et al.*, 2004). Une fois présent dans le génome, la machinerie cellulaire est utilisée afin d'exprimer la totalité des protéines virales, tout comme n'importe quel autre gène. Pour les rétrovirus simples, l'activation de cette transcription est possible grâce aux promoteurs présents dans les régions non-codantes bordant le génome proviral aussi appelées LTR. Les rétrovirus complexes, quant à eux, orchestrent plus précisément l'expression de leurs protéines en produisant des trans-

activateurs qui contrôlent la transcription d'une panoplie de gènes. Finalement, il y a traduction du matériel nécessaire à l'assemblage des particules virales et bourgeonnement des virions à l'extérieur de la cellule.

L'étape d'intégration de certains rétrovirus peut être à l'origine de plusieurs problèmes métaboliques puisqu'une fois dans le génome de la cellule hôte, les séquences promotrice présentes dans les LTR viraux activent également la transcription de gènes cellulaires bordant les sites d'intégrations. Plusieurs recherches ont mis en évidence l'activation d'oncogènes ou l'expression de micro ARN par des intégrations rétrovirales menant au développement de cancers (Denicourt et al., 1999; Denicourt et al., 2003; Landais et al., 2007; Uren et al., 2005).

1.2 Les rétrovirus HTLV

Les rétrovirus HTLV (*Human T-cell Lymphotropic Virus*) et STLV (*Simian T-cell Lymphotropic Virus*) ont été désignés comme étant des PTLV (*Primate T-cell Lymphotropic Virus*), puisqu'ils sont issus d'un ancêtre commun. Ils partagent une organisation génomique semblable, en plus de présenter des propriétés virologiques et épidémiologiques similaires.

1.2.1 Origine et distribution des différents membres

Des études phylogéniques ont révélé que HTLV-1 a fort possiblement été transmis du singe vers l'homme, il y a $50\,000 \pm 10\,000$ ans (Van Dooren et al., 2001), mais ce n'est qu'en 1980 que le groupe de recherche du Dr Gallo a détecté et isolé pour une première fois le rétrovirus humain HTLV-1, à partir d'une population de lymphocyte T d'un patient atteint de lymphome cutané des cellules T (*cutaneous T-cell lymphoma*) (Poiesz et al., 1980). Rapidement, un bon nombre d'études ont mis en évidence la constitution génomique et protéomique de ce virus présent au Japon, en Afrique centrale, en Amérique du sud et dans le bassin des Caraïbes. On estime aujourd'hui qu'il y a entre dix et vingt millions d'individus porteurs de ce rétrovirus à travers le monde. Le second rétrovirus de cette famille, HTLV-2, a été découvert deux années plus tard par le même groupe de recherche, soit en 1982 à partir

d'échantillon sanguin d'un patient souffrant de leucémie lympho-histiocytaire (*hairy T-cell leukemia*) (Kalyanaraman et al., 1982). Aucune donnée n'est disponible concernant le nombre d'individus infectés à travers le monde, mais on parle d'endémie en Afrique centrale, chez les populations amérindiennes d'Amérique du nord, centrale et du sud. De plus, ce rétrovirus est particulièrement présent chez les utilisateurs de drogues par voie intraveineuse en Europe et aux États-Unis.

Récemment, les troisième et quatrième membres de la famille des HTLV, soient HTLV-3 et HTLV-4 ont été isolés chez des Camerounais. L'attribution de ces rétrovirus à la famille HTLV s'est basée sur une étude comportant un test d'immunofluorescence en présence d'anticorps anti-HTLV-1 et -2, d'analyses de réactions de polymérase en chaîne (PCR) faisant appel à des amorces hautement conservées dans les gènes *pol* et *tax* des rétrovirus humains HTLV-1, -2 et simien STLV-3, en plus d'analyses phylogéniques de ces mêmes gènes, comparativement au rétrovirus simien STLV-3 (Calattini et al., 2005; Wolfe et al., 2005).

1.2.2 Pathogénicité et transmission

Le rétrovirus HTLV-1 a été le premier pathogène identifié à avoir été associé à la formation de leucémie, en plus d'être la première preuve irréfutable qu'une infection virale peut être à l'origine du développement de cancer chez l'humain. En effet, ce dernier a la capacité d'immortaliser les cellules lymphocytaires CD4+, pouvant mener au développement de graves pathologies (Burmeister, 2001).

La majorité des individus HTLV-1-positifs vont rester asymptomatiques tout au long de leur vie, mais un certain pourcentage d'entre eux va tout de même développer des maladies agressives et mortelles, telles que des ATL (*adult T-cell leukemia*) ou des dérégulations du système immunitaire, telles des neurodégénérescences aussi nommées HAM/TSP (*HTLV-1-associated myelopathy/tropical spastic paraparesis*). Le risque de développer un jour une ATL pour les porteurs est d'environ 6,6% pour les hommes et de 2,1% pour les femmes (Arisawa et al., 2000). Cette leucémie des cellules T de l'adulte se développe généralement après vingt à trente ans de latence et peut se présenter sous quatre formes. Dans un premier temps, il y a les porteurs chroniques, qui ont 5% ou plus de

lymphocytes T anormaux tout en conservant un niveau normal de lymphocytes sains dans le sang. Ces derniers, peuvent également présenter des lésions cutanées ou pulmonaires. Ensuite, la leucémie chronique est caractérisée par une augmentation des lymphocytes dans le sang (lymphocytose), une hypertrophie des ganglions lymphatiques (lymphoadénopathie) impliquant le foie, la rate, la peau et les poumons, en plus de présenter plus de 5% de lymphocytes T anormaux dans le sang. Le type lymphome se définit par une lymphoadénopathie avec ou sans lésion extra-ganglionnaire mais en ayant toutefois uniquement 1% ou moins de lymphocytes T anormaux. Finalement, le type aigu inclut les patients souffrant d'ATL et de lésions mais qui n'entrent dans aucune des trois catégories précédentes. Ces derniers présentent généralement des lymphocytes à noyaux polylobés typiques (figure 1.3) (Shimoyama, 1991).

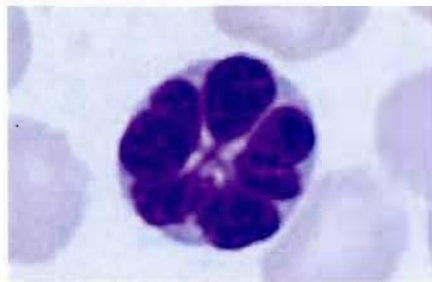


Figure 1.3 : Cellule sanguine d'un patient atteint d'ATL, présentant un lymphocyte T à noyau multilobulé. Image tirée de Matsuoka, 2005.

D'un autre côté, la paraparésie spastique tropicale ou myélopathie associée à HTLV-1 touche moins de 1% des individus infectés. Cette dernière se caractérise par une démyélinisation débutant au niveau de la moelle épinière, provoquant des modifications de la motricité volontaire, de la coordination des mouvements et du tonus musculaire (spasticité) des membres inférieurs. De plus, on note une réponse exagérée du système parasympathique à diverses excitations, se manifestant entre autres par de l'hypertension, de la rougeur à la figure et au thorax (hyperréflexie), en plus d'incontinence urinaire et de légères pertes de sensibilité périphérique. HTLV-1 peut se transmettre de plusieurs façons : lors de rapports sexuels, au moment de l'allaitement maternel (le transport de lymphocytes T infectés à

travers la barrière placentaire est très rare) et par le sang contaminé (Burmeister, 2001; Gallo, 2005).

En ce qui concerne le rétrovirus HTLV-2, aucune étude ne permet de démontrer avec certitude qu'il est l'agent causal d'une quelconque maladie, mais une infection par ce dernier a quelques fois été associée à des leucémies lympho-histiocytaires et à quelques cas de désordres neurologiques pouvant s'apparenter aux HAM/TSP. Le sang est le meilleur véhicule de transmission, mais les rapports sexuels et l'allaitement maternel permettent également la propagation du virus (Feuer and Green, 2005).

Pour l'instant, aucune pathologie n'est associée à une infection par HTLV-3 et HTLV-4, probablement en raison du trop faible nombre de cas répertoriés et de la très longue latence des virus. De plus amples études portant sur ces nouveaux rétrovirus devraient aider à comprendre les mécanismes d'infection possibles (Calattini et al., 2005; Wolfe et al., 2005).

1.2.3 Structure

Les virions de HTLV-1 et HTLV-2 ont des diamètres approximatifs de 100 nm et contiennent leur matériel génétique en deux copies sous forme d'ARN monocaténaire, de polarité positive, ayant une extrémité 5' coiffée par une 7-méthylguanosine triphosphate (GTP), aussi appelée séquence «cap» et une queue de poly A en 3', laquelle consiste en une chaîne d'adénosines. L'ARN est contenu dans la capsid, en présence des enzymes transcriptase inverse (RT), intégrase et protéase. La nucléocapsid (p15) protège le génome en formant une structure compacte lorsque liée à l'ARN. La capsid (p24) est contenue à l'intérieur de la matrice (p19), laquelle est enveloppée d'une bicouche phospholipidique provenant de la cellule hôte dans laquelle sont insérés les complexes glycoprotéiques d'enveloppe de surface (SU) et transmembranaire (TM) d'origine virale, soit respectivement les protéines virales gp46 et gp21, chez le rétrovirus HTLV-1 (figure 1.4) (Poiesz et al., 2003).

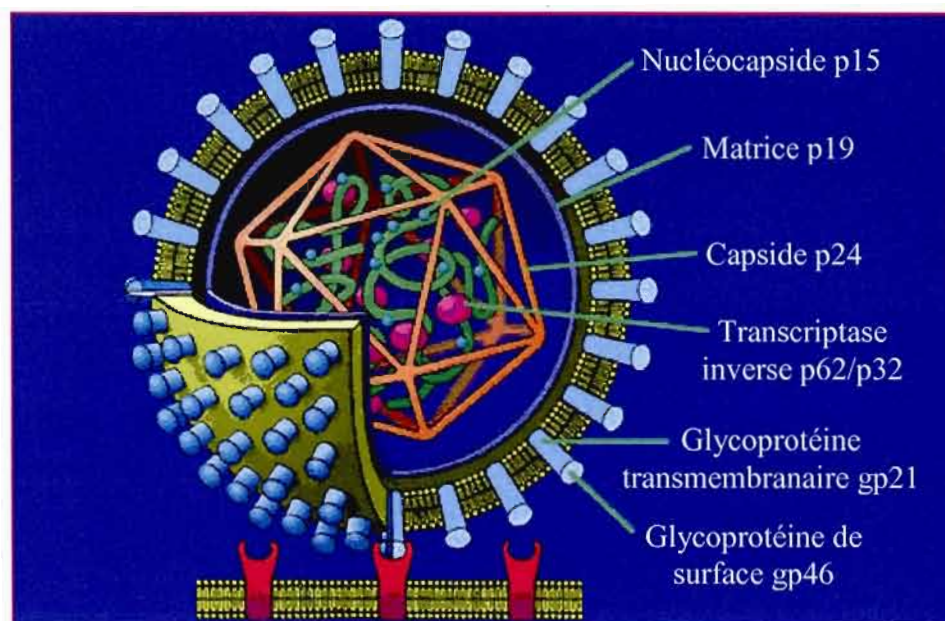


Figure 1.4 : Structure du rétrovirus HTLV-1. Le matériel génétique sous forme d'ARN est contenu dans la capsidie en présence de la nucléocapsidie et des enzymes transcriptase inverse, protéase et intégrase. La capsidie est entourée de la matrice, laquelle est recouverte d'une membrane de phospholipides provenant de la cellule hôte. Les glycoprotéines gp21 et gp46 sont présentes à la surface des virions de HTLV-1. Image tirée et modifiée de : http://fr.wikipedia.org/wiki/Virus_de_l'immunod%C3%A9ficiency_humaine.

1.2.4 Mécanisme d'infection

Le contact entre les cellules semble important pour une infection par le rétrovirus HTLV-1 puisque peu de particules virales sont produites et que très peu d'entre elles sont infectieuses (une sur 10^5 à 10^6) (Derse et al., 2001; Fan et al., 1992). De ce fait, une infection par le rétrovirus HTLV-1 nécessite la formation d'une synapse virologique, permettant la transmission virale. Cette dernière est initiée par la rencontre de deux cellules T, l'une infectée par HTLV-1 et l'autre non. Quarante minutes plus tard, une forte polarisation a lieu au contact entre les cellules, caractérisée par une abondance des protéines virales Gag (p19 et p15). La protéine de l'enveloppe qui est exprimée à la surface des cellules infectées joue potentiellement un rôle dans l'initiation de cette polarisation. Deux heures plus tard, l'ARN

génomique d'HTLV-1 se transfère vers la seconde cellule, de même que les protéines Gag (Igakura et al., 2003).

Il semble que le transporteur ubiquitaire de glucose GLUT-1 soit à l'origine de la liaison entre les protéines de l'enveloppe des rétrovirus HTLV-1 et HTLV-2 et leurs cellules cibles (Manel et al., 2003). Il a été rapporté qu'une activation des lymphocytes T par le TCR (*T cell receptor*) engendre une expression accrue de GLUT-1, résultant en une augmentation de la capture de glucose par ces cellules, en plus d'une liaison plus importante des protéines d'enveloppes de HTLV-1 et -2 aux lymphocytes T CD4+ et CD8+ (Kinet et al., 2007). La présence de HSPG (*heparan sulfate proteoglycans*) est également essentielle à l'attachement et à l'entrée des virions de HTLV-1, tandis que ce composé n'a pas d'effets similaires chez HTLV-2 (Jones et al., 2008). Les cellules lymphocytaires T CD4+ semblent toutefois être la première cible du rétrovirus HTLV-1, mais des cellules infectées CD8+ et des lymphocytes B ont également été identifiées *in vivo*. D'un autre côté, le rétrovirus HTLV-2 infecte préférentiellement des cellules lymphocytaires T CD8+.

Récemment, il a été démontré que HTLV-1 a également la capacité d'infecter des cellules dendritiques myéloïdes et plasmacytoïdes, lesquelles peuvent à leur tour infecter les lymphocytes T CD4+. L'infection des cellules CD4+ par les cellules dendritiques nécessite la présence de NP-1 (*neuropilin-1*) et de HSPG. Cette découverte représente la première preuve convaincante qu'une infection par le rétrovirus HTLV-1 ne requiert pas nécessairement un contact cellule-cellule et suggère que les cellules dendritiques jouent potentiellement un rôle important dans la transmission, la propagation et la persistance de HTLV-1 *in vivo* (Jones et al., 2008).

1.2.5 Organisation génomique

Les rétrovirus de la famille des HTLV ont un génome d'environ 9 kb qui code pour les quatre gènes essentiels à la réplication virale, soit les gènes *gag*, *pro*, *pol* et *env*. La région *gag* code pour les protéines de la matrice (p19), du capsid (p24), et du nucléocapsid (p15), le gène *pro* code pour une protéase, la région *pol* est responsable de la réplication et de l'intégration du virus grâce aux enzymes transcriptase inverse et intégrase, tandis que la séquence *env* code pour des protéines de surface et transmembranaire. Deux régions non

codantes appelées LTR sont situées aux extrémités et contiennent chacune les régions U3, R et U5 (figure 1.5). La région R contient les signaux d'initiation de la transcription et de polyadénylation, la partie U5 a été rapporté comme favorisant une polyadénylation efficace des transcrits viraux, en plus de contenir des sites de liaison pour l'enzyme intégrase (Böhnlein et al., 1989 ; Bouziane et al., 1996) et la région U3 constitue une région promotrice et activatrice de la transcription virale.

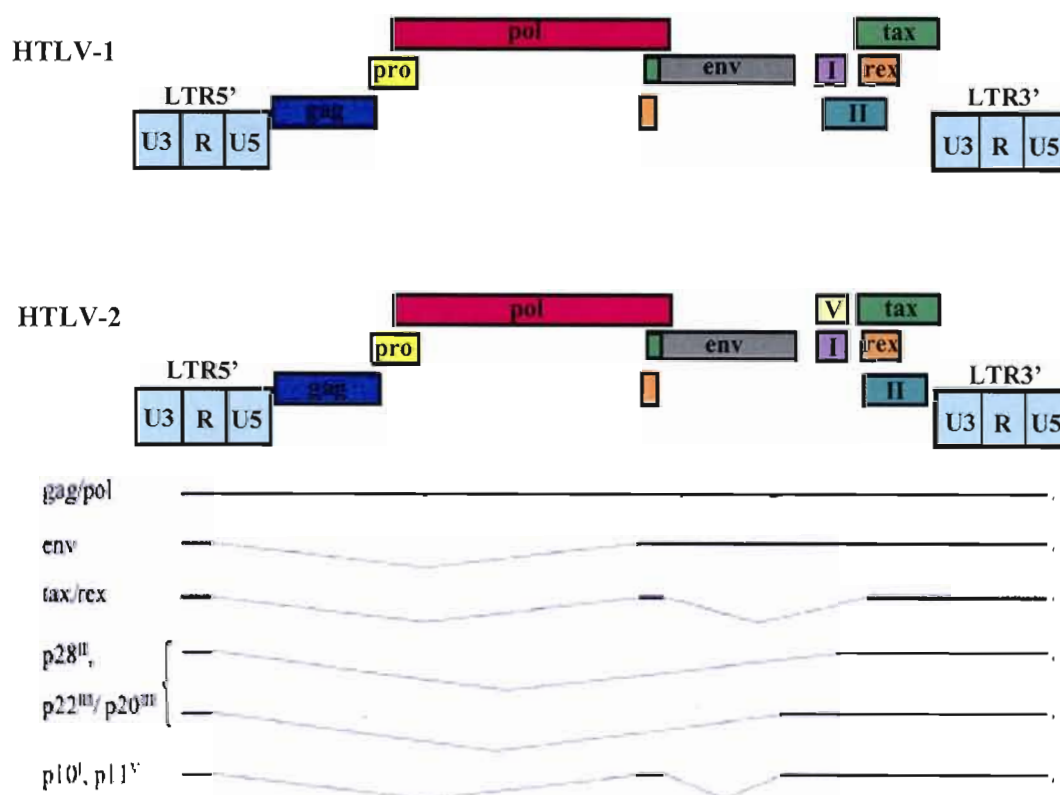


Figure 1.5 : Génome et structure des transcrits des rétrovirus HTLV-1 et HTLV-2. Les génomes de ces rétrovirus contiennent les gènes *gag*, *pro*, *pol*, *env*, *rex*, *tax*, les ORF I et II ainsi que l'ORF V chez HTLV-2, encadrés par deux LTR constitués des régions U3, R et U5. Image modifiée de Feuer et Green, 2005.

En plus de ces protéines essentielles, les rétrovirus HTLV-1, -2, -3 et -4 contiennent le matériel génétique pour produire des protéines accessoires et régulatrices. Ces dernières sont essentiellement situées dans les cadres de lecture ouverts de la région pX, situés dans la

portion 3' du génome et produisent les protéines Tax et Rex (figure 1.5). La protéine Tax joue le rôle de transactivateur de la transcription initiée au promoteur de la région U3 du LTR5'.

D'autre part, Rex régule l'expression virale au niveau post-transcriptionnel en stabilisant et augmentant l'export nucléaire d'ARN messenger (ARNm) partiellement ou non épissé dans le cytoplasme (Younis et al., 2005). Les rétrovirus HTLV-1 et HTLV-2 codent également pour des protéines accessoires qui sont impliquées dans le processus d'infection, de transmission et de persistance virale, soit les ORF I et II codant pour les protéines p12/p27 et p13/p30 pour HTLV-1 et les ORF I, II et V, produisant les protéines p10, p28 et p11 pour HTLV-2 (Bartoe et al., 2000; Cockerell et al., 1996; Collins et al., 1998; Feuer and Green, 2005; Silverman et al., 2004).

1.2.5.1 Tax et Rex

Les protéines accessoires Tax et Rex jouent un rôle important dans le cycle de réplication virale puisqu'elles régulent l'expression des différentes protéines au niveau transcriptionnel et post-transcriptionnel. Immédiatement après l'infection d'une cellule T, la transcription des gènes viraux dépend uniquement des interactions entre les facteurs de transcription cellulaires et le LTR5' du rétrovirus. Il y a d'abord expression et exportation des transcrits complètement épissés, codant pour les protéines Tax et Rex vers le cytoplasme. La traduction de Tax est largement favorisée au dépend de celle de Rex, ce qui mène à l'expression d'une grande quantité de Tax comparativement à Rex. Tax est une protéine de 40 kDa principalement localisée dans le noyau. Sa présence en grande quantité dans la cellule a pour effet d'augmenter la transcription du promoteur viral en se liant aux facteurs de transcription cellulaires liés aux séquences CRE (*cAMP response element*) présentes dans les TRE (*Tax responsive elements*) de la région U3 du LTR viral. Les protéines qui se fixent à la séquence CRE contiennent des motifs *leucine zipper* au même titre que les protéines de la famille CREB/ATF. La formation du complexe tertiaire liant Tax, CREB et le promoteur rétroviral semble essentiel à la transcription virale (Feuer and Green, 2005). Cette augmentation de la transcription a pour effet d'augmenter la traduction de Tax et Rex. Cette dernière est une phosphoprotéine nucléolaire constituée de 27 kDa, qui accroît l'exportation

vers le cytoplasme des ARNm viraux génomiques (*gag/pol/pro*) et mono-épissés (*env*) menant à la traduction de protéines virales de structure, d'assemblage et de bourgeonnement. Cette exportation est directement dépendante d'une structure secondaire que forme l'ARNm nommée élément de réponse à Rex (RexRE). En effet, le génome de HTLV-1 contient une séquence comprise à l'intérieur des régions U3 et R de son LTR3' qui, lorsque transcrite, forme quatre tiges boucles constituant le RexRE. C'est précisément l'interaction entre la protéine Rex et le RexRE des ARNm qui permettent leur export vers le cytoplasme (Kashanchi and Brady, 2005).

Le processus par lequel une infection par les rétrovirus HTLV-1 et -2 mène à la transformation des lymphocytes T implique assurément la protéine *Tax* (Feuer and Green, 2005). Le mécanisme précis n'est pas encore bien défini, mais des études proposent que *Tax* engendre des dérégulations au niveau transcriptionnel et post-transcriptionnel chez les cellules lymphocytaires infectées. La différence de pathogénicité chez les patients infectés par HTLV-1 ou -2 pourrait être attribuable à la variation dans la constitution protéique de *Tax*, qui partage une homologie de 78% chez ces deux rétrovirus (Feuer and Green, 2005). *Tax* module l'expression des gènes et des facteurs de transcription cellulaires comme des facteurs de réponse au sérum (SRF) et NF- κ B. Par exemple, *Tax* active les facteurs de transcription de la famille NF- κ B au niveau du cytoplasme et du noyau, ce qui engendre l'expression de gènes dont le promoteur contient ce motif. Il s'ensuit la transcription de gènes comme l'interleukine-2 (IL-2), le récepteur α de l'interleukine-2, l'interleukine-3 (IL-3) ainsi qu'une variété de gènes y compris des proto-oncogènes comme c-Fos (Feuer and Green, 2005). La protéine *Tax* interagit également avec plusieurs protéines de la cellule hôte, ce qui dérégule le cycle cellulaire, la stabilité et la réparation de l'ADN. En effet, des études ont démontré que *Tax* a l'habileté d'induire le passage du cycle cellulaire de la phase G1 à S en plus de perturber la transition de la phase G2 à M, ce qui explique la présence de cellules lymphocytaires contenant plusieurs noyaux chez les patients leucémiques (Jin et al., 1998). La protéine *Tax* intervient également dans la stabilité de l'ADN puisqu'elle réprime la transcription de l'enzyme qui ajoute des séquences nucléotidiques répétées aux extrémités des chromosome: la télomérase ou hTERT (Gabet et al., 2003). Dans le même ordre d'idée, *Tax* altère l'intégrité de l'ADN en interférant avec différents mécanismes de réparation de l'ADN tels BER (*base excision repair*), NER (*nucleotide excision repair*), MMR (*mismatch*

repair) et DSB_R (*DNA double-strand break repair*). En effet, l'expression de plusieurs facteurs impliqués dans ces mécanismes, incluant l'ADN polymérase β et PCNA (*proliferating cell nuclear antigen*) est modulée par l'action de Tax, résultant en une augmentation de la fréquence des mutations cellulaires (Lemoine et al., 2000; Marriott and Semmes, 2005; Miyake et al., 1999).

1.3 La transcription

Chez les cellules eucaryotes, la transcription nécessite l'interaction entre une ARN polymérase et le promoteur du gène à transcrire. Il existe trois types d'ARN polymérase (I, II et III) ayant chacune une fonction bien précise. Les ARN polymérases I et III permettent la transcription d'ARN ribosomique, de transfert ainsi que plusieurs petits ARN. D'un autre côté, l'ARN polymérase II est responsable de l'expression des précurseurs d'ARN messagers, ce qui représente la grande majorité de la transcription cellulaire. Le processus de transcription requiert également des protéines de liaison à l'ADN (TFII pour « *Transcription Factor of RNA polymérase II* ») qui se lient d'abord à la matrice d'ADN pour ensuite interagir avec l'ARN polymérase puisque aucune de ces polymérases n'est capable de s'y lier adéquatement (Alberts et al., 2002).

Ainsi, la transcription de l'ARN messager nécessite d'abord la liaison de TFIID à la boîte TATA présente sur le promoteur, plus particulièrement environ 25 nucléotides en aval du site d'initiation de la transcription. TFIID est une protéine multimérique qui contient entre autres la TBP (*TATA binding protein*) qui va directement se lier au petit sillon de la boîte TATA, mais qui est grandement stabilisée par l'arrivée de TFIIA (figure 1.6) (Lee and Keller, 1991; Starr and Hawley, 1991).

Ensuite, puisque le complexe TFIID/TFIIA ne peut pas lier l'ARN polymérase II correctement, TFIID lie TFIIB. Cette interaction constitue l'étape limitante de la transcription de plusieurs gènes et une fois en place, l'ARN polymérase vient s'y lier. Avant ou après sa liaison avec TFIIB, l'ARN polymérase II s'associe avec TFIIF et TFIIIE (Buratowski et al., 1991; Conaway and Conaway, 1991). TFIIF a la capacité de dérouler l'ADN alors que TFIIIE est une ATPase dépendante de l'ADN, probablement responsable de la génération d'énergie nécessaire à la transcription (Bunick et al., 1982; Sawadogo and Roeder, 1984). Finalement,

l'ARN polymérase qui est lié à TFIID est libérée par l'activité de la protéine kinase de TFIIF, qui phosphoryle les résidus du domaine C-terminal de l'ARN polymérase II, permettant l'initiation de la transcription (Akoulitchev et al., 1995).

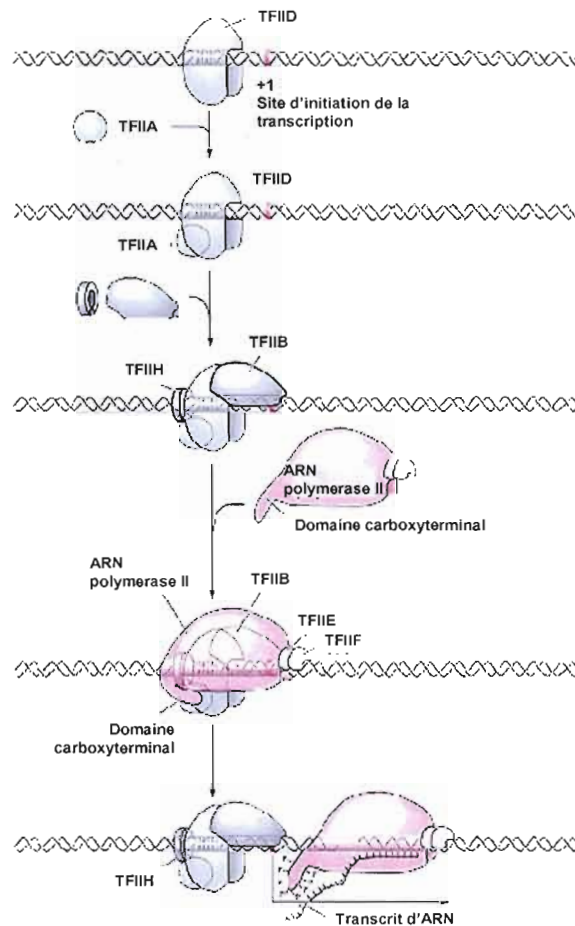


Figure 1.6 : Représentation du complexe d'initiation de la transcription chez les organismes eucaryotes. Dans un premier temps, le complexe TFIID se lie à la boîte TATA *via* la sous unité TBP. Ensuite, l'arrivée de TFIIA stabilise TFIID jusqu'à l'arrivée de TFIIB et TFIIF. Le complexe d'ARN polymérase II, TFIIE et TFIIF se positionne sur TFIIB, tandis que le domaine C-terminal de l'ARN polymérase II se lie à TFIID. Le domaine C-terminal de l'ARN polymérase II est alors phosphorylé par TFIIF et libéré par TFIID, ce qui initie la transcription. Image modifiée de Gilbert, 2005.

Il existe également plusieurs gènes qui ne contiennent pas de boîte TATA dans leur séquences promotrices et qui sont tout de même transcrits par l'ARN polymérase II. Dans ces cas, d'autres séquences vont permettre la formation du complexe de pré-initiation de la transcription (PIC) telles que des *initiators* et des DPE (*downstream promoter element*). La formation de ces complexes peut être associée à plus d'un site d'initiation de la transcription et différentes protéines telles que Sp1 vont lier le promoteur, permettant une liaison adéquate de TFIID, ce qui initie la cascade de la formation du complexe transcriptionnel (Gilbert, 2005).

1.3.1 Polyadénylation

Lorsque l'ARN polymérase II termine la transcription d'un gène et qu'elle rencontre des signaux de polyadénylation, elle est capable d'ajouter une série d'adénines à la fin du pré-ARNm afin de le stabiliser. Le signal consensus de polyadénylation correspond à la séquence nucléotidique AAUAAA. Le clivage du transcrit a généralement lieu de dix à trente nucléotides en aval de ce signal et moins de trente nucléotides séparent le site de clivage d'une région riche en nucléotides G et U. Ainsi, les protéines CstF (*cleavage stimulation factor*) et CPSF (*cleavage and polyadenylation specificity factor*) sont transportées par l'ARN polymérase II et se lient au pré-ARNm lorsqu'elles en ont la possibilité. L'hexamère AAUAAA lie la protéine CPSF, tandis que la région riche en G et en U interagit avec CstF. D'autres protéines sont également recrutées afin de former le complexe qui clive d'abord l'ARN en un site bien précis. Ensuite, une polymérase nommée PAP (*poly A polymerase*) ajoute un à un les résidus désoxyadénosine jusqu'à concurrence d'environ deux cents. Lorsque la queue de poly A est complétée, des protéines nommées PABP (*poly-A-binding proteins*) s'y lient et par un mécanisme encore méconnu, déterminent la longueur de la chaîne d'adénines. Ces protéines accompagnent l'ARNm lors de son export au cytoplasme et aident à diriger la synthèse protéique (Alberts et al., 2002).

1.3.2 Épissage

Plusieurs transcrits eucaryotes, y compris les transcrits viraux, subissent des modifications post-transcriptionnelles dans le noyau avant leur exportation vers le cytoplasme. L'épissage est l'une de ces modifications et elle permet d'augmenter le potentiel

codant du génome de la cellule en joignant différentes combinaisons d'exons, engendrant ainsi la production d'une multitude de protéines. En effet, immédiatement après leur transcription, les transcrits jusque-là nommés pré-ARNm contiennent des exons et des introns. Ainsi, l'intron non codant est extrait *via* deux réactions de trans-estérification résultant en la jonction des deux exons (figure 1.7).

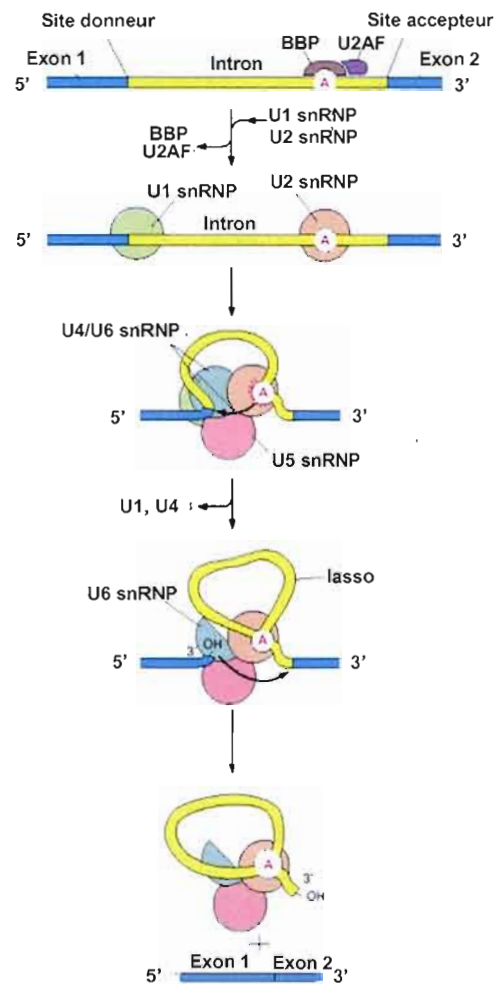


Figure 1.7 : Schéma des étapes d'épissage. Le point de branchement est reconnu par BBP (*branch point binding protein*) et U2AF. Ensuite U2 snRNP lie le point de branchement, ce qui éjecte BBP et U2AF, tandis que U1 snRNP lie le site donneur d'épissage. L'arrivée de U4, U6 et U5 snRNP forme le spliceosome et une série de réarrangements engendre l'éjection de U1 et U4 snRNP. Le groupement hydroxyl du site donneur d'épissage réagit ensuite avec le site accepteur, joignant les deux exons et libérant l'intron sous forme de lasso. Image modifiée de Alberts *et al.*, 2002.

Le mécanisme d'épissage implique le spliceosome constitué de cinq courtes molécules d'ARN (U1, U2, U4, U5 et U6 nommés snARN) complexées à au moins sept sous-unités de protéines formant les snRNP (*small nuclear ribonucleoprotein*). Le mécanisme d'épissage débute par l'assemblage du spliceosome sur le pré-ARNm. La reconnaissance des sites donneurs et accepteurs d'épissage ainsi que du point de branchement est due à un pairage de bases entre l'ARN et les snARN. Ensuite, une désoxyadénosine présente dans l'intron attaque le site donneur d'épissage et clive la liaison entre le sucre et le groupement phosphate de l'ARN. L'extrémité 5' de l'intron est alors liée de façon covalente à l'adénine, créant une boucle dans l'ARN semblable à un lasso. L'hydroxyle nucléophile du site donneur d'épissage réagit ensuite avec le site accepteur, joignant les deux exons et libérant de ce fait l'intron sous forme de lasso. L'intron sera alors dégradé dans le noyau alors que l'ARNm sera exporté vers le cytoplasme en circulant à travers les pores nucléaires de la membrane nucléaire (Alberts et al., 2002).

1.4 Transcription antisens

1.4.1 **Chez les organismes procaryotes et eucaryotes**

La transcription antisens était depuis longtemps uniquement décrite chez les organismes procaryotes, lesquels font appel à ce mécanisme pour réguler négativement la traduction des transcrits sens en plus d'être impliqué dans le contrôle de certaines fonctions biologiques telles que la transposition, la réplication de plasmides et la conjugaison (Vanhee-Brossollet and Vaquero, 1998). Il y a quelques années, la transcription antisens de gènes complémentaires aux régions codantes de protéines ayant des fonctions biologiques très diversifiées a été décrite chez des insectes, des oiseaux et des mammifères. Des études suggèrent que la transcription antisens a une fonction de régulation négative des transcrits sens au même titre que chez les organismes procaryotes, mais leur fonction réelle n'est pas encore établie (Briquet et al., 2001; Miller, 1988; Vanhee-Brossollet and Vaquero, 1998).

1.4.2 Chez les rétrovirus

La transcription antisens a pour la première fois été proposée chez le rétrovirus humain VIH-1 suite à une analyse bioinformatique permettant d'identifier un cadre de lecture ouvert conservé chez une douzaine de clones (Miller, 1988). La région codante identifiée se localise vis-à-vis celle de la protéine de l'enveloppe et engendrerait la production d'une protéine constituée d'environ 190 acides aminés à caractère apparemment hydrophobe, ce qui permet d'émettre l'hypothèse que cette dernière serait associée à une structure membranaire. Considérant que la protéine d'enveloppe du VIH-1 est reconnue pour contenir de nombreuses mutations, la conservation d'un cadre de lecture ouvert sur le brin complémentaire de cette dernière devient difficilement attribuable à un phénomène de hasard et ajoute un argument en faveur de l'importance de cette transcription antisens lors de l'évolution de ce virus. Cette transcription antisens a été démontrée et la localisation de la protéine codée par ce transcrit a fait l'objet d'une publication (Briquet and Vaquero, 2002; Bukrinsky and Etkin, 1990; Landry et al., 2007; Michael et al., 1994). La transcription antisens a également été étudiée chez le rétrovirus HTLV-1; il s'agit de la caractérisation la plus exhaustive et la plus convaincante d'une transcription antisens chez un rétrovirus. En effet, HTLV-1 code pour une protéine antisens nommée HBZ pour *HTLV-1 bZIP factor* et on lui attribue un rôle dans la régulation de la transcription (Gaudray et al., 2002).

1.4.3 Chez HTLV-1

La transcription antisens chez HTLV-1 a pour la première fois été proposée en 1989 par le groupe de Brunck, lesquels ont détecté un transcrit antisens avec une sonde d'ARN simple brin lors d'une hybridation de type Northern blot (Larocca et al., 1989). Par la suite, plus d'une dizaine d'années ont passé avant que l'équipe de recherche du Dr Mesnard effectue une expérience de double hybride à partir d'une banque d'ADNc provenant de cellules chroniquement infectées par HTLV-1, soit la lignée cellulaire MT2 (Gaudray et al., 2002). L'expérience consistait en la recherche de partenaires d'interaction du facteur de transcription CREB-2 *via* le motif bZIP qu'il contient. Ainsi, ce groupe a identifié une nouvelle protéine qu'il a attribuée au brin complémentaire du génome de HTLV-1. La séquence de cette dernière n'avait jusqu'alors jamais été rapportée dans la banque de donnée

« GenBank » et compte tenu qu'elle contenait les caractéristiques d'un facteur de transcription possédant un domaine bZIP, le nom « *HTLV-1 bZIP factor* » ou HBZ lui a été attribué. En effet, la partie C-terminale détient les caractéristiques d'un motif *leucine zipper*. Plus précisément, les résidus 164 à 192 contiennent une isoleucine ainsi que quatre leucines, chacune séparée par six acides aminés, lesquels sont précédés d'une région riche en arginines et en lysines (acides aminés 140 à 163) pouvant être nécessaire à la liaison de cette protéine à l'ADN. À cet instant, la caractérisation du transcrit antisens n'avait pas été complétée. Ainsi, le cadre de lecture ouvert situé vis-à-vis les nucléotides 7292 à 6666 du brin sens du génome proviral de HTLV-1 était ciblé comme étant la région codante de cette protéine virale nouvellement identifiée constituée de 209 acides aminés. C'est en 2006 que trois publications signées par les groupes de Matsuoka, Kamihira et Barbeau ont rapporté que le transcrit antisens de HTLV-1 est épissé modifiant ainsi la séquence protéique jusqu'alors prédite (Cavanagh et al., 2006; Murata et al., 2006; Satou et al., 2006). En effet, le phénomène d'épissage résulte en l'ajout d'un premier exon (contenant treize nucléotides codants) lequel est localisé dans le LTR3' et en la perte des vingt-deux premiers nucléotides du cadre de lecture ouvert préalablement identifié (figure 1.8). De ce fait, il y a production de la protéine HBZ-SP1, cette fois-ci constituée de 206 acides aminés. Depuis la découverte de cette protéine, plusieurs groupes de recherche se sont intéressés à cette transcription antisens, ce qui a permis la caractérisation tant du transcrit, du promoteur que de la protéine.

													EXON 1	EXON 2		
													M A A S	G L F	HBZSP1	
													ATG GCG GCC TCA G...	GG CTG TTT		
*	T	S	R	V	R	Q	S	V	E	S	R	L	S	L	G L F	HBZSP2
TGA	ACA	AGC	AGG	GTC	AGG	CAA	AGC	GTG	GAG	AGC	CGG	CTG	AGT	CTA	G...	
													M V N F V S V	G L F	HBZ NON-ÉPISSÉ	
													ATG GTT AAC TTT GTA TCT GTA G	GG CTG TTT		

Figure 1.8 : Représentation des séquences nucléotidiques et protéiques autour du site d'épissage des différents isoformes de HBZ. On observe HBZ-SP1, SP2 et HBZ non-épisé. Image modifié de Cavanagh *et al.*, 2006.

1.4.3.1 Structure du transcrit

La caractérisation du transcrit antisens de HTLV-1 a été publiée quatre ans après la découverte de la protéine issue de cette transcription, soit en 2006. Ainsi, l'initiation de la transcription a lieu dans le LTR3', plus précisément dans les régions R et U5 sur une zone couvrant approximativement 182 nucléotides. On sait également que ce transcrit est polyadénylé et le signal consensus de polyadénylation est situé à 1450 nucléotides en aval du codon stop de l'ORF d'HBZ (Cavanagh et al., 2006). De plus, l'épissage alternatif de ce dernier engendre l'expression de deux transcrits, soit SP1 qui est plus abondant et SP2. Dans ce cas d'épissage alternatif, on retrouve deux sites donneurs différents qui se lient au même site accepteur présent dans le cadre de lecture ouvert de la protéine HBZ. Notons que le codon initiateur de HBZ-SP1 se localise dans le premier exon à la position 8689 (selon la numérotation sens) et que la forme SP2 ne contient pas de codon initiateur classique (AUG). Il a toutefois été proposé qu'un autre codon tel que GUG ou CUG initie la traduction de la forme SP2, considérant qu'il a été rapporté que ce type de processus a lieu pour d'autres protéines de HTLV-1 (Cavanagh et al., 2006; Corcelette et al., 2000; Murata et al., 2006). De récentes études ont également fait état d'un second type de transcrit, lequel ne serait pas épissé et initierait plutôt à différents endroits à proximité du nucléotide 1461 (selon la numérotation antisens) du génome de HTLV-1. Le codon initiateur du cadre de lecture ouvert de ce transcrit se localise au nucléotide 7298 (selon la numérotation sens), soit à la fin de l'intron de HBZ-SP1 (figure 1.8). L'activité promotrice de ces deux types de transcrits est différente, quoique tous deux stimulées en la présence de Tax (Yoshida et al., 2008).

1.4.3.2 « HTLV-1 bZIP factor »

La découverte d'une nouvelle protéine issue d'une transcription antisens est une autre preuve de l'efficacité des rétrovirus à produire leurs différentes protéines à partir d'un génome de taille limitée. Jusqu'alors, le processus entourant le développement de leucémies chez le faible pourcentage d'individus infectés restait à définir, mais voici qu'un nouveau partenaire viral entre dans l'équation : HBZ.

1.4.3.2.1 Structure moléculaire

L'isoforme HBZ-SP1 est la forme la plus abondante et il s'agit d'une protéine nucléaire de 31 kDa composée de 206 acides aminés, qui est localisée principalement dans le noyau. Cette protéine contient dans l'ordre : un domaine d'activation de la transcription en *N*-terminal (AD), deux régions composées d'acides aminés basiques impliqués dans le transport au noyau (BR2 et BR1), un domaine de liaison à l'ADN (DBD) et un motif *leucine zipper* en *C*-terminal (ZIP), lequel permet de former des complexes avec des facteurs de transcription (figure 1.9) (Hivin et al., 2005).

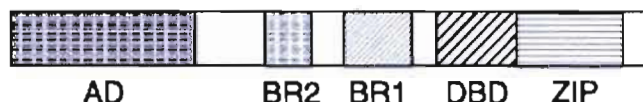


Figure 1.9 : Structure moléculaire de la protéine HBZ. Elle contient dans l'ordre, un domaine d'activation de la transcription (AD), deux régions composées d'acides aminés basiques (BR2 et BR1), un domaine de liaison à l'ADN (DBD) et un motif *leucine zipper* (ZIP). Image tirée de Hivin *et al.*, 2005.

Le motif *leucine zipper* est souvent retrouvé chez des facteurs de transcription cellulaires tels que CREB, CREB-2, Fos et Jun, permettant une dimérisation avec d'autres protéines contenant ce motif. La création d'un *leucine zipper* nécessite une interaction entre deux structures semblables, chacune constituée d'une courte hélice α contenant un résidu leucine à chaque sept acides aminés. Ainsi, chaque leucine est en contact direct avec une autre leucine présente sur la seconde hélice α avec laquelle elle interagit. Ce type de motif est généralement accompagné d'un domaine de liaison à l'ADN, constitué d'acides aminés basiques, permettant sa liaison *via* des ponts hydrogènes (Figure 1.10). Cette structure particulière est à l'origine de son implication dans la régulation de l'expression génique (Landschulz et al., 1988).

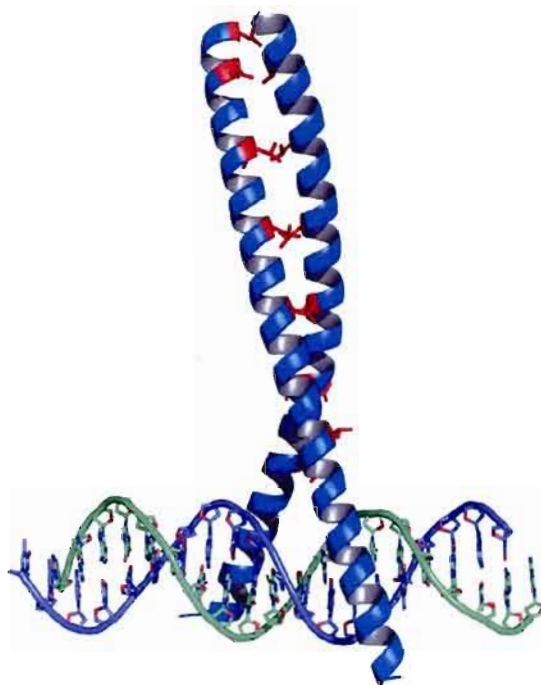


Figure 1.10 : Structure tridimensionnelle d'un *leucine zipper* interagissant avec l'ADN. Le *leucine zipper* est constitué de deux hélices α (bleu) liées *via* les résidus leucines (rouge). Image de Thomas Splettstoesser, tirée de : http://commons.wikimedia.org/wiki/Image:Leucine_zipper.png.

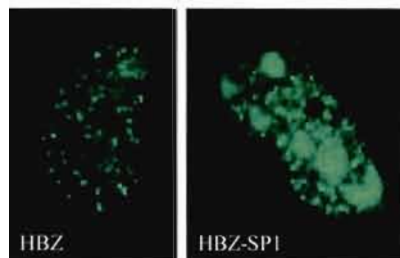
1.4.3.2.2 Activité promotrice

L'étude du promoteur de la transcription antisens a été présentée très récemment et les recherches font la distinction entre l'activité promotrice de la forme non-épissée et la forme SP1. Comme mentionné précédemment, l'initiation de la transcription de la forme épissée de HBZ-SP1 a lieu dans les régions U5 et R du LTR3' et il n'y a pas de boîte TATA en amont. En ce qui concerne la forme non-épissée (usHBZ), les sites d'initiation de la transcription ont été situés dans les environs du nucléotide 1461 (numérotation du transcrit antisens) et encore une fois, aucune boîte TATA n'a été identifiée en amont. D'un autre côté, des séquences d'initiation putatives et des éléments promoteurs consensus ont été identifiés en aval des sites d'initiation de la transcription des deux formes d'HBZ. L'activité promotrice de la forme HBZ-SP1 s'est avérée supérieure à celle de la forme usHBZ, ce qui concorde avec une plus forte

expression du transcrit de l'isoforme d'HBZ épissée détectée chez les cellules infectées (Usui et al., 2008). Il semble également que la transcription de la forme épissée soit fortement dépendante de l'un des trois sites de liaison Sp1 localisé en amont. De plus, la protéine virale Tax stimule la transcription de usHBZ *via* les trois TRE (*Tax responsive element*), alors que l'effet est moins important sur la forme HBZ-SP1 (Yoshida et al., 2008). De plus, la localisation des sites d'intégration de HTLV-1 dans le génome de la cellule hôte semble être impliquée dans la stimulation de la transcription par Tax. En effet, cette activation est inversement proportionnelle à la transcription basale antisens. Ces résultats impliquent qu'une intégration provirale dans une région transcriptionnellement peu active subira une plus forte stimulation transcriptionnelle antisens par la protéine virale Tax (Landry et al., 2008).

1.4.3.2.3 Localisation cellulaire

La première étude de localisation cellulaire a été achevée en 2002 à partir de l'isoforme non-épissé de la protéine. La fusion d'une étiquette fluorescente verte communément appelée GFP (*green fluorescent protein*) en *N*-terminal de la protéine a permis une détection d'HBZ sous forme granulaire dans le noyau (Gaudray et al., 2002). Considérant cette localisation, de plus amples études ont permis d'identifier plus précisément les signaux de localisation nucléaires (NLS). Pour plusieurs facteurs de transcription contenant un domaine bZIP, les acides aminés basiques contenus dans leurs domaines de liaison à l'ADN sont suffisants pour permettre leur exportation vers le noyau. Ainsi, des expériences de mutation par délétions ont permis l'identification de trois régions impliquées dans le transport d'HBZ vers le noyau. Il s'agit en fait des deux régions basiques BR1 et BR2, en plus du domaine de liaison à l'ADN (DBD). Par contre, une combinaison d'uniquement deux de ces trois domaines est nécessaire à son transport (Hivin et al., 2005). D'un autre côté, le patron de distribution de l'isoforme HBZ-SP1 diverge légèrement de la précédente. En effet, des expériences semblables ont démontré que HBZ-SP1 se localise également dans le noyau, mais cette fois-ci en colocalisant avec les nucléoles (Figure 1.11). De tels résultats suggèrent qu'une modification de quelques acides aminés en *N*-terminal (suppression de 7AA et ajout de 4AA) est suffisante pour possiblement permettre une interaction avec des partenaires cellulaires différents, ce qui engendre une localisation nucléaire altérée.



Figures 1.11 : Localisation cellulaire des différents isoformes de HBZ dans les cellules COS-7 par microscopie confocale. La distribution granulaire de HBZ diffère de celle de HBZ-SP1 qui colocalise avec les nucléoles. Images tirées de Murata *et al.*, 2006.

1.4.3.2.4 Fonctions

La caractérisation des différents domaines constituant la structure de HBZ laisse présager un rôle pouvant s'apparenter aux facteurs de transcription contenant des motifs semblables. Ainsi, des études de fonctionnalité ont permis l'identification de différents partenaires d'interaction résultant en une régulation de la transcription ayant des conséquences diverses sur l'expression des protéines virales et cellulaires.

Dans un premier temps, il a été démontré que HBZ interagit avec le facteur de transcription CREB-2 *via* son motif bZIP résultant en une régulation négative de la transcription virale (Gaudray et al., 2002). En effet, il a préalablement été mis en évidence que le facteur de transcription CREB-2 a la capacité de se lier à la protéine virale Tax pour former un complexe stable constitué des protéines CREB-2/Tax/CBP capable de lier les TRE présents sur le LTR de HTLV-1 (Gachon et al., 2000). De ce fait, la liaison d'HBZ avec CREB-2 l'empêche ainsi de lier les CRE et les TRE, résultant en une régulation négative de l'expression de la totalité des protéines virales. Dans le même ordre d'idées, HBZ a également la capacité d'interagir avec le facteur de transcription CREB, encore une fois par son *leucine zipper*, ce qui a pour conséquence d'empêcher la liaison de ce dernier au CRE de HTLV-1, résultant en une diminution de la transcription virale (Lemasson et al., 2007). L'inhibition de la transcription virale dépendante de Tax est également attribuable à deux motifs composés des acides aminés LXXLL (plus précisément VDGLL et LDGLL) situés en *N*-terminal de la protéine HBZ. Ces derniers interagissent directement avec les coactivateurs p300 et CBP par l'intermédiaire de leur domaine KIX. Cette interaction a pour conséquence de prévenir la liaison de Tax à

p300/CBP, ce qui diminue la transcription virale (Clerc et al., 2008).

Ensuite, considérant que les cellules T transformées par HTLV-1 expriment un haut niveau d'ARNm codant pour des facteurs de transcription de la famille AP-1, tels que c-Jun, JunB, JunD et c-Fos, il devient important de vérifier si cette nouvelle protéine virale a la capacité d'interagir avec ces facteurs de transcription apparemment impliqués dans la dérégulation des lymphocytes T chez les individus HTLV-1 positifs (Fujii et al., 1991). Ainsi, il a été rapporté que HBZ est capable de dimériser avec les facteurs de transcription JunB et c-Jun diminuant ainsi leur capacité de lier l'ADN. Les dimères HBZ/JunB sont séquestrés dans des corps nucléaires inhibant ainsi son activité transcriptionnelle (Hivin et al., 2007). Il a été proposé qu'en présence d'HBZ, le facteur de transcription c-Jun, qui forme habituellement un homodimère, soit plus susceptible de se lier à HBZ formant ainsi un hétérodimère incapable de se lier aux sites AP-1 présents sur le LTR de HTLV-1 (Basbous et al., 2003). Le fait que les hétérodimères HBZ/c-Jun et HBZ/JunB soient incapables de lier l'ADN pourrait s'expliquer par l'absence de certains résidus importants du domaine de liaison à l'ADN chez la protéine HBZ, empêchant ainsi leur attachement au site AP-1. Cette hypothèse a d'ailleurs été vérifiée et il s'avère que la séquence EQERRE (acides aminés 134 à 139 de la forme non-épissée, ou 131 à 136 de la forme HBZ-SP1) localisée en amont du domaine de liaison à l'ADN soit responsable de l'effet répresseur d'HBZ sur l'activité de c-Jun (Hivin et al., 2006). De plus, des études ont démontré que la diminution de la transcription dépendante de c-Jun par HBZ implique également une réduction de la stabilité de c-Jun, la dirigeant vers une dégradation par le protéasome (Matsumoto et al., 2005). D'un autre côté, l'association de HBZ avec JunB a pour effet d'augmenter légèrement mais significativement la transcription initiée par le promoteur de la collagénase, lequel contient un élément de liaison AP-1. Cette activation de la transcription peut s'expliquer par le fait que l'hétérodimère HBZ/JunB, qui possède un très faible potentiel de liaison de l'ADN, soit plus stable que l'homodimère JunB favorisant la transcription des gènes dépendant de ce facteur transcriptionnel. Une autre étude a aussi mis en évidence la séquestration de JunB dans des corps nucléaires (Hivin et al., 2007).

De plus, l'interaction entre HBZ et JunD a également fait l'objet d'études et l'hétérodimère HBZ/JunD active la transcription initiée par le promoteur de la collagénase de manière dose dépendante, contrairement au dimère HBZ/c-Jun. De plus amples études ont permis d'identifier que le domaine de liaison à l'ADN localisé en *N*-terminal de la protéine

HBZ est responsable de cette stimulation de la transcription dépendante de JunD. De plus, lorsque le domaine de liaison à l'ADN (DBD) d'HBZ est absent, il y a diminution de la transcription, laissant croire qu'en absence de cette région, la dimérisation reste possible par le *leucine zipper*. Certains ont émis l'hypothèse que ces résultats peuvent également s'expliquer par le fait que la liaison de HBZ à JunD préviendrait la liaison de co-répresseurs (Thebault et al., 2004). Il semble que le même motif EQERRE, préalablement démontré pour être responsable de la perte d'activité de c-Jun, soit dans ce cas impliqué dans l'activation de la transcription médiée par JunD (Hivin et al., 2006). Cette activation de la transcription médiée par le facteur transcriptionnel JunD implique nécessairement la modulation d'une quantité de gènes cellulaires ayant peut-être un rôle à jouer dans le mécanisme de transformation cellulaire ou dans certaines étapes du développement de la maladie chez les individus infectés. Ainsi, HBZ augmente la transcription de la sous-unité catalytique de la télomérase hTERT *via* les sites AP-1 présents sur son promoteur (Kuhlmann et al., 2007). Considérant qu'il a été rapporté qu'il existe une forte corrélation entre l'évolution de l'ATL et l'activité de la télomérase, cela laisse présager le rôle potentiel de HBZ dans le développement des leucémies (Uchida et al., 1999).

La protéine HBZ n'est pas indispensable à l'infection et à l'immortalisation des lymphocytes T puisque la prolifération cellulaire est la même chez des PBMC (*Peripheral Blood Mononuclear Cell*) infectées par HTLV-1 sauvage ou HTLV-1 mutant pour HBZ. Par contre, HBZ détient un rôle dans l'augmentation de l'infection et de la persistance virale. Une étude effectuée *in vivo* chez des lapins infectés par HTLV-1 sauvage ou mutant pour HBZ a permis de confirmer l'importance de cette protéine au début d'une infection. En effet, uniquement deux semaines après l'infection, les lapins infectés par les virus mutants présentent une plus faible charge virale en plus d'une plus faible réponse antigénique (Arnold et al., 2006).

Finalement, une étude qui met en lumière le rôle du transcrit antisens a démontré son caractère prolifératif. Des expériences visant à éteindre l'expression de cette protéine chez des cellules leucémiques en faisant appel à un vecteur lentiviral recombinant exprimant une courte tige boucle d'ARN dirigée contre HBZ ont montré que l'expression d'HBZ est nécessaire à la prolifération des cellules ATL. Étonnamment, la mutation du codon initiateur de HBZ (méthionine) par une leucine n'a aucun impact négatif sur l'effet prolifératif de ce dernier,

suggérant que ces observations sont attribuables au transcrit plutôt qu'à la production de la protéine en tant que telle. Cette hypothèse a été confirmée puisque l'effet prolifératif n'est plus observable lorsque la totalité des codons d'HBZ subissent des mutations silencieuses (mutations nucléotidiques n'affectant pas la production de l'acide aminé) qui peuvent généralement altérer la structure du transcrit. Ces résultats supposent que l'ARN codant pour la protéine HBZ puisse former une structure tridimensionnelle pouvant agir directement sur la prolifération cellulaire (Satou et al., 2006).

Bref, les fonctions d'HBZ sont multiples et cette nouvelle protéine virale semble détenir un rôle dans la modulation de la transcription cellulaire et virale dépendante de Tax *via* les motifs leucine zipper et LXXLL qu'elle contient. HBZ favorise également l'infection et permet la persistance virale, tandis que la structure tridimensionnelle que forme le transcrit contribue à la prolifération cellulaire (figure 1.12).

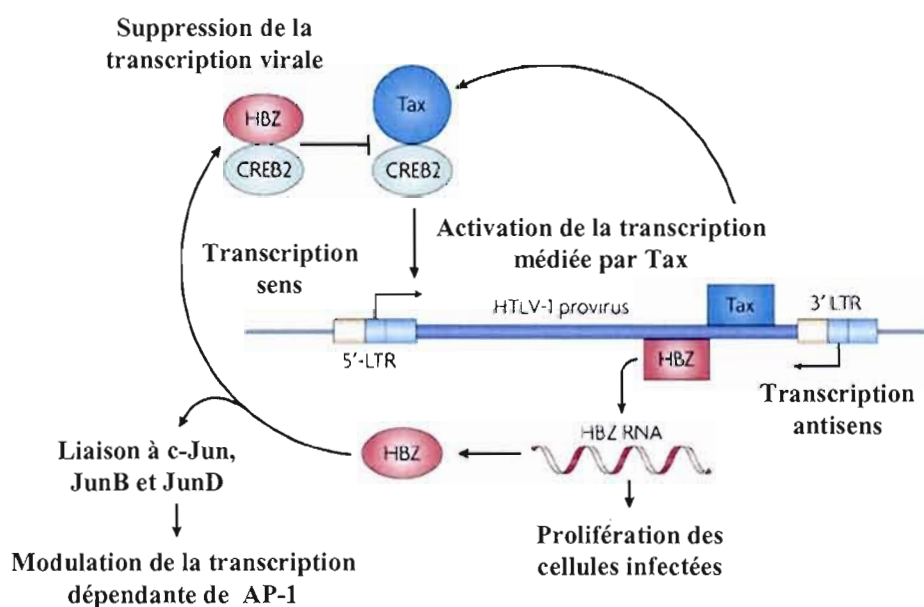


Figure 1.12 : Schéma des différentes actions de HBZ et conséquences sur la transcription. HBZ module la transcription dépendante de AP-1, inhibe la transcription virale dépendante de Tax, tandis que l'ARN permet la prolifération cellulaire. Image modifiée de Matsuoka et Jeang, 2007.

1.4.3.2.5 HBZ et ATL

Le mécanisme responsable du développement de l'ATL n'est pas encore complètement compris mais plusieurs données permettent de faire un lien entre l'expression de la protéine HBZ et le développement de cette maladie. Dans un premier temps, HBZ module l'expression de plusieurs gènes cellulaires tels que hTERT, par l'entremise de son interaction avec différents facteurs de transcription (Kuhlmann et al., 2007). Ensuite, il a été démontré que HBZ stimule la prolifération cellulaire en plus de favoriser la croissance et l'infiltration des tumeurs chez des souris (Arnold et al., 2008; Satou et al., 2006). Des études ont également mis en évidence que les cellules ATL expriment constamment la protéine HBZ alors que l'expression de Tax n'est pas constante (Usui et al., 2008). Ces observations concordent avec le fait que les intégrations rétrovirales des cellules ATL présentent des LTR3' intacts alors que les LTR5' sont souvent absents ou non fonctionnels (Miyazaki et al., 2007). Ainsi, ces informations suggèrent que l'expression d'HBZ pourrait supporter la survie des cellules infectées et potentiellement l'oncogenèse.

1.5 Hypothèse de travail

L'étude de la transcription antisens chez les différents rétrovirus de la famille HTLV est devenue essentielle à la compréhension des possibles mécanismes de transformation cellulaire associés à une infection rétrovirale. Il a en effet été rapporté que chez les individus souffrant d'ATL, les intégrations rétrovirales sont caractérisées par des LTR5' absents ou fortement méthylés alors que les LTR3' restent intacts. Ainsi, la totalité des protéines virales issues d'une transcription sens ne sont pas exprimées, tandis que la protéine HBZ est produite (Miyazaki et al., 2007). De ce fait, la dérégulation des lymphocytes T chez un certain pourcentage d'individus HTLV-1 positifs pourrait être dûe ou supportée par l'expression d'HBZ.

D'un autre côté, le rétrovirus HTLV-2 n'engendre pas le même genre de pathogénicité chez les individus infectés et une infection par les rétrovirus HTLV-3 et HTLV-4 n'est, pour l'instant, pas associée au développement d'une maladie. La raison d'une telle ou possible divergence entre les symptômes associés à une infection par ces rétrovirus

reste à définir. Compte tenu qu'aucune étude n'a été menée à ce jour concernant la transcription antisens des rétrovirus HTLV-2, -3 et -4, il devient essentiel de clarifier si ces derniers expriment une protéine jusque-là ignorée. L'hypothèse de cette recherche réside dans le fait que ces rétrovirus pourraient également exprimer un transcrit antisens lequel coderait pour une nouvelle protéine virale potentiellement différente d'HBZ. L'identification de nouvelles protéines virales devient essentielle à la compréhension de mécanismes d'infection et de persistance virale.

Ainsi, les objectifs de la recherche ont été la détection et la caractérisation de la transcription antisens chez les rétrovirus humain HTLV-2, -3 et -4. La structure ainsi que les extrémités des transcrits antisens ont été caractérisés. L'identification des cadres de lectures ouverts a permis d'effectuer des études de localisation cellulaire en plus de vérifier l'impact de ces nouvelles protéines sur la transcription virale dépendante de Tax. Finalement des études de promoteur ont permis l'identification de composés activant la transcription antisens. Il se pourrait qu'une différence de structure et/ou de fonctionnalité de ces protéines antisens puisse expliquer les divergences de pathogénicité des virus faisant l'objet de cette étude.

CHAPITRE II

ARTICLE 1

2.1 Contribution des auteurs

L'article a été rédigé par l'auteur du présent mémoire, Marilène Halin, avec les précieux conseils de son directeur de recherche, Benoît Barbeau, ainsi que des collaborateurs Renaud Mahieux et Jean-Michel Mesnard. Marilène Halin a également effectuée des RT-PCR, les expériences de RACE, le Northern blot et les alignements de séquences, en plus de participer activement au montage des figures. Estelle Douceron a fait des RT-PCR, les essais d'inhibition de la transcription virale et les Westerns blots. Isabelle Clerc a réalisé les expériences de localisation cellulaire. Chloée Journo et Nga Ling Ko ont fait des RT-PCR. Sébastien Landry a effectué des RT-PCR et il a participé à la planification de certaines expériences. Edward Murphy a fourni les échantillons d'ARN de patients infectés. Jean-Michel Mesnard a participé à l'écriture de l'article; il a de même offert une expertise en microscopie confocale. Benoit Barbeau et Renaud Mahieux ont mis sur pied le projet et ils ont contribué à l'écriture de l'article et au montage des figures.

La mise en forme de l'article correspond aux exigences de présentation des mémoires de l'Université du Québec à Montréal. L'article a été soumis au journal *Blood*.

2.2 Résumé

Les rétrovirus HTLV-1 et HTLV-2 infectent les lymphocytes T humains. Il a été démontré que le virus HTLV-1 induit la formation de leucémie chez un certain pourcentage de patients infectés. Par ailleurs, le rétrovirus HTLV-2 n'a pas encore été associé au développement de maladie hématologique. Le brin complémentaire du génome du rétrovirus HTLV-1 code pour une protéine nommée HBZ, qui pourrait potentiellement être impliquée dans le développement de leucémie chez les patients infectés. Dans cet article, nous démontrons que le brin complémentaire du génome de HTLV-2 code également pour une protéine que nous avons nommée APH-2 (*Antisense Protein of HTLV-2*). L'ARN_m de APH-2 est épissé, polyadénylé et la transcription est initiée dans le LTR3' à différentes positions. Le transcrit a été détecté dans différentes lignées cellulaires chroniquement infectées par HTLV-2 et également chez un individu infecté. La protéine APH-2 est constituée de 183 acide aminés et est localisée dans le noyau des cellules transfectées. Malgré l'absence d'un motif *leucine zipper* conservé, APH-2 réprime l'activation de la transcription virale induite par la protéine Tax.

Nos résultats démontrent l'expression d'une nouvelle protéine antisens chez le rétrovirus HTLV-2, laquelle pourrait être impliquée dans le développement de certaines pathologies, telle la lymphocytose fréquemment observée chez les patients infectés par HTLV-2.

2.3 Human T-cell Leukemia Virus Type 2 produces a spliced antisense transcript encoding a protein that lacks a bZIP domain but inhibits Tax2-mediated transcription

Marilène Halin^{1*}, Estelle Douceron^{2*}, Isabelle Clerc³, Chloé Journo², Nga Ling Ko², Sébastien Landry¹, Edward Murphy⁴, Jean-Michel Mesnard³, Benoît Barbeau^{1†}, Renaud Mahieux^{2†}

¹Département des Sciences Biologiques et Centre de recherche BioMed, Université du Québec à Montréal, 2080 Saint-Urbain Montréal, Québec, Canada H2X 3X8. ²Unité d'Epidémiologie et Physiopathologie des Virus Oncogènes, CNRS URA 3015, Département de Virologie, Institut Pasteur, 28 rue du Dr Roux, 75015 Paris, France. ³Centre d'études d'agents Pathogènes et Biotechnologies pour la Santé (CPBS), CNRS/UM1/UM2 UMR 5236, Montpellier, France ⁴Blood Systems Research Institute and University of California, San Francisco, San Francisco, USA.

* and † These authors contributed equally

Correspondence to Renaud Mahieux, Unité d'Epidémiologie et Physiopathologie des Virus Oncogènes, CNRS URA 3015, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris cedex 15, France. Tel: (33)-1-45-68-89-06, Fax: (33)-1-40-61-34-65, E-mail: renaud.mahieux@pasteur.fr or to Benoît Barbeau, Département des Sciences Biologiques, Université du Québec à Montréal, 2080 Saint-Urbain Montréal, Québec, Canada H2X 3X8. Tel: 514 987-3000 ext.4576, Fax: (514) 987-4647, E-mail: barbeau.benoit@uqam.ca

2.3.1 Abstract

Human T-cell Leukemia Virus type 1 and type 2 retroviruses infect T-lymphocytes. While HTLV-1 infection causes leukemia, HTLV-2 has not been demonstrated to be the agent of a hematological malignant disease. The minus strand of the HTLV-1 genome encodes HBZ, a protein which could play a role in the development of leukemia in infected patients. Herein, we demonstrate that the complementary strand of the HTLV-2 genome also encodes a protein that we named APH-2 (Antisense Protein of HTLV-2). *APH-2* mRNA is spliced, polyadenylated and initiates in the 3'LTR at different positions. This transcript was detected in different tested HTLV-2-infected cell lines as well as in one HTLV-2 carrier. The APH-2 protein is 183 amino acid long and is localized in the nucleus of transfected cells. Despite the lack of a bZIP domain consensus sequence, APH-2 represses Tax2-mediated transcription. Our results demonstrate the existence of an antisense strand-encoded protein in HTLV-2, which could represent an important player in the development of disorders such as the lymphocytosis that is frequently observed in HTLV-2 patients.

2.3.2 Introduction

HTLV-1 and HTLV-2 retroviruses belong to the Primate T Lymphotropic Virus family (Verdonck et al., 2007). They share a number of features including a similar genomic organization, similar cellular receptor complexes, the presence of a pX region that encodes a viral Tax transactivator protein and several accessory proteins (Feuer and Green, 2005). Interestingly however, while HTLV-1 is the causal agent of Adult T-Cell Leukemia/Lymphoma (ATLL), a fatal disease that occurs in 1-5% of infected individuals, HTLV-2 infection is only associated with some rare cases of a neurological disease resembling Tropical Spastic Paraparesis/HTLV-1-Associated Myelopathy (Roucoux and Murphy, 2004). In recent years, several distinctions have been reported between Tax1 and Tax2 in terms of their transactivating abilities as well as their intracellular distribution (Mahieux et al., 2000; Meertens et al., 2004a; Ross et al., 1997; Sakamoto et al., 1992; Semmes et al., 1996; Tanaka et al., 1996; Tripp et al., 2003; Van et al., 2001). These differences might account in part for the distinctive outcomes observed in HTLV-1 vs HTLV-2-infected patients. However, the poor abundance of *tax1* mRNA in ATLL patients represents an observation which is difficult to reconcile with the suggested primary role played by the viral protein in the maintenance of a transformed stage.

HBZ (HTLV-1 bZIP factor) is a protein, which was surprisingly shown to be encoded by the complementary strand of the HTLV-1 genome. The discovery of this protein provided a promising explanation for the mechanism of ATLL development (Gaudray et al., 2002). HBZ transcripts initiate in the 3'LTR and are either unspliced or spliced, thereby producing two isoforms, HBZ and HBZ-SP1 (also named HBZ-SI) respectively (Cavanagh et al., 2006; Murata et al., 2006; Satou et al., 2006). The HBZ-SP1 transcript is present in all HTLV-1-infected cells, i.e. primary cells from healthy carriers and ATLL patients as well as in cell lines, and its level is 4-fold higher than that of unspliced HBZ mRNA in cell lines (Usui et al., 2008). Furthermore, HBZ-SP1 expression was also shown to correlate with the HTLV-1 proviral load (Usui et al., 2008). A recent study also demonstrated that, even if *HBZ* mRNA levels are lower (10 fold) than *tax/rex* mRNA levels, they are however much more abundant (10-1000 fold) than p12, p27/p12, p30 and p13 accessory gene transcripts (Li and Green, 2007).

The HBZ protein is 209 amino-acid long (Gaudray et al., 2002), while HBZ-SP1 encodes for a 206 amino acid isoform (Cavanagh et al., 2006; Murata et al., 2006). These two proteins only differ in their *N*-terminal sequences, while both contain a bZIP (basic leucine zipper) domain at their COOH-terminus. HBZ also possesses three different Nuclear Localization Signals (NLS), two of which are necessary for its localization to the nucleus (Hivin et al., 2005). In this cellular organelle, HBZ forms nuclear bodies that are different from PML bodies or splicing factor compartments (Hivin et al., 2005), while HBZ-SP1 is found in the nucleolus (Hivin et al., 2007; Murata et al., 2006).

In vitro, both isoforms down-regulate Tax-mediated viral transcription, although a recent report has suggested that HBZ-SP1 is a stronger repressor of Tax-mediated transcription than the unspliced HBZ isoform (Yoshida et al., 2008). Tax inhibition is linked to the ability of both isoforms to interact either with the CREB and CREB-2 transcription factors and/or to sequester the CBP/p300 transcription co-activators (Clerc et al., 2008; Lemasson et al., 2007). HBZ is also capable of forming dimers with c-Jun, JunD or JunB, which consequently alters their transcriptional activity (Hivin et al., 2006; Hivin et al., 2007; Kuhlmann et al., 2007; Thebault et al., 2004). *In vivo*, experiments using either HBZ-SP1 transgenic mice or a rabbit infected with wild-type or HBZ-deficient HTLV-1 viruses led a number of investigators to suggest that HBZ was responsible for (1) the proliferation of CD4⁺ lymphocytes cells and (2) the viral persistence *in vivo*, but not for cellular transformation (Arnold et al., 2006; Arnold et al., 2008; Satou et al., 2006).

Interestingly, recent data obtained from a cohort of HTLV-2-infected carriers demonstrated that the total lymphocyte count is increased by 7-10% in these individuals when compared to non-infected blood donors (Ed Murphy, unpublished data). We therefore hypothesized that HTLV-2 complementary strand might encode a protein that would behave like HBZ and which could participate in the cell proliferation *in vivo* but which could also down-regulate the viral expression.

In this study, we demonstrate that the complementary strand of the HTLV-2 genome encodes a protein that we have tentatively named APH-2 (Antisense Protein of HTLV-2). *APH-2* mRNA was detected in different tested HTLV-2-infected cell lines regardless of the subtype as well as in one HTLV-2 carrier. *APH-2* mRNA is spliced, polyadenylated and initiates in the 3'LTR at different positions. The APH-2

protein is 183 amino-acid long with an apparent molecular weight of ≈ 25 kDa. In transfected cells, APH-2 localizes to the nucleus but not to the nucleolus unlike HBZ-SP1. APH-2 lacks a bZIP domain but nonetheless represses Tax2-mediated transcription. Whether APH-2 is responsible for the lymphocytosis that is observed in HTLV-2 patients remains to be determined.

2.3.3 Material and methods

2.3.3.1 Cell lines

The HTLV-2-transformed cell lines C19, BJAB/Gu and MO, as well as HTLV-1 transformed C91/PL or uninfected Jurkat and CEM cells were grown in RPMI medium supplemented with 10% fetal bovine serum and antibiotics (penicillin 100 U/ml, streptomycin 100 µg/ml). 293T and HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with fetal bovine serum (10%) and antibiotics (penicillin 100 U/ml, streptomycin 100 µg/ml). All cell lines were maintained at 37°C in 5% CO₂.

2.3.3.2 HTLV-2 carrier samples

Peripheral blood mononuclear cells (PBMC) (5 million/sample) were obtained from 5 different HTLV-2-infected subjects, who were previously enrolled in the HOST (HTLV-1 Outcomes Study). For each subject, the proviral load had previously been determined and ranged from 10³ to 10⁴ per 10⁶ PBMC (data not shown).

2.3.3.3 Plasmids

The HTLV-2 proviral DNA clone pH6neo (Chen et al., 1983; Shimotohno et al., 1985) was kindly provided by Dr. P. Green. The pHTLV-2ΔEcoRI construct was generated by cloning the 4269 bp EcoRI fragment (nucleotide 4683 to 8952) in pBlueScript KS (see Figure 2.1A). The APH-2 cDNA fragment including the complete coding region was cloned into the pEGFP-C2 (Clontech) and pcDNA3.1-Myc-His C (Invitrogen) expression vectors using the EcoRI and EcoRI/HindIII restriction sites, respectively.

2.3.3.4 Transfections

293T cells were transfected either using the calcium phosphate protocol or with the Polyfect reagent (Qiagen) while HeLa cells were transfected with the Effectene reagent (Qiagen) as previously described (Fortin et al., 1997; Meertens et al., 2004a).

2.3.3.5 RT-PCR, 5' and 3' RACE analyses

Total RNA was extracted using the TRIzol reagent (Invitrogen) as recommended by the manufacturer, and treated twice with the DNaseI RNase-free DNA set (Qiagen) to avoid any carry-over of the proviral plasmid as previously described (Chevalier et al., 2008).

PolyA⁺ RNA was purified from total RNA using the Oligotex mRNA Mini Kit (Qiagen) according to manufacturer's instructions. RT-PCR analyses were conducted using the RT primer 24-2 (5'-CTATACACTCCAAGTCTGATGCC-3') (Figure 2.1A). Briefly, total RNA (5µg) was mixed with 0.5µM of the specific RT primer. The RNA:RT primer mix was heated at 70°C for 5 min, cooled on ice for 5 min and incubated for 2 h at 42°C in the presence of 1X AMV reaction buffer (1 mM dNTPs, 10U of SUPERase-In RNase inhibitor (Ambion) and 15 U AMV reverse transcriptase (USB)). cDNAs were then amplified in the presence of 1.25 U Taq DNA polymerase, 1X ThermoPol buffer, 20µM dNTPs, 1.5µM of each primer using a Tgradient thermocycler (Biometra). Primers added to the PCR reactions were the reverse primer 21-2 (5'-CTCTCCCGGCGCTTTTGATCC-3') and forward primers 19-1 (5'-CGTATGTTGGCCCCGGGAC-3'), 22-1 (5'-CTCGACCTGAGAGGAGACTTAC-3') or 21-1 (5'-GGGGTAATCGTCAGAGCCTTG-3') (Figure 2.1A). PCR conditions were as followed: denaturation at 94°C for 5 min followed by 35 cycles (94°C 1 min, 60°C 1 min, 72°C 1 min) and a final extension at 72°C for 5 min. RT-PCR amplifications were controlled for DNA contamination (RNA samples with no RT step). For amplification of the total encoding APh-2 cDNA, PCR amplification was conducted on RNA from transfected 293T cells using the 22-1 forward and 22-6 reverse (5'-TGTATCCTCCGCCAAATCCAGG-3') primers (positioned at nt 293-314 and 2381-2402 of the antisense strand, respectively). The

amplified product was then cloned in pBlueScript KS and further used either for generating the APH-2 expression vectors or as a probe for Northern-blot analyses.

Both the 5' and 3' extremities of the antisense transcript were PCR amplified from isolated polyA⁺ RNA using the SMART RACE cDNA Amplification kit (Clontech) and the FirstChoice RLM-RACE kit (Ambion), respectively, according to manufacturer's instructions. For the 5' end, the cDNA was synthesized with a modified oligo (dT) and ligated to a supplied anchor at their 5' end. Two subsequent PCR rounds were conducted with the 5'RACE outer and inner primers along with the 24-2 and 21-2 transcript-specific reverse primers. For the 3' end, cDNA synthesis was performed in the presence of the supplied 3'RACE adapter complementary to the polyA tail. PCR amplification was conducted in the presence of the 3'RACE outer and inner primers and the forward 21-3 primer (5'-GAGAAGTTGGTGGAGGCTGAG-3') derived from the sequence downstream of the HTLV-2 antisense APH-2 ORF (nt 3224-3244 of the antisense strand). PCR products were cloned in pBlueScript KS and sequenced.

2.3.3.6 RT-PCR analyses in infected cell lines

Total RNA was extracted from MO and C19 HTLV-2 cell lines, or from uncultured PBMC obtained from five HOST patients with the TRIzol reagent and treated twice with DNase. 5µg of total RNA were then used for the RT-PCR experiment with the SuperScript Reverse Transcriptase II and RNase Out ribonuclease inhibitor kit (Invitrogen). During the RT step, the APH-2-specific 24-2 primer was used. The cDNA was then amplified with a semi-nested PCR using the 19-1/21-2 primer set for the first PCR followed by the 22-1/21-2 primer set for the second PCR. The GAPDH primers were described elsewhere (Bialek et al., 2003).

2.3.3.7 Northern blot analyses

Four micrograms of polyA⁺ RNA isolated from 293T cells were denaturated in formaldehyde/formamide for 10 min at 70°C, chilled on ice and separated on a 1.5% agarose

gel containing MOPS (20mM) and formaldehyde (17%). After migration, the gel was then transferred onto a Hybond-XL nylon membrane (Amersham Biosciences). The membrane was UV cross-linked for 3 min and hybridized with a ³²P-labeled APH-2 encoding cDNA in a hybridization buffer (50% formamide, 4X SSPE, 5X Denhardt's, 0.5% SDS, 10% Dextran sulfate and 0.1 mg/ml of denaturated salmon sperm DNA). The hybridized blot was washed twice in 2X SSC for 15 min at 42°C and the membrane was exposed to BioRad Imaging screen K and revealed with a PhosphorImager and the Quantity One software (BioRad Molecular Image FX).

2.3.3.8 Generation of APH2-antibodies

Based on the deduced APH-2 amino acid sequence, two synthetic oligopeptides representing separate domains of the protein were synthesized (Eurogentec). The peptides were selected from regions in the sequence found to be hydrophilic. Oligopeptide A, amino-acids n°21-41 and oligopeptide B, amino-acids n°45-58 of the APH-2 sequence were coupled to keyhole limpet hemocyanin (KLH) and injected to specific pathogen-free New Zealand white rabbits. After several booster injections, the serum (SPY557) was tested for the presence of anti-APH-2 antibodies. As a negative control, a pre-immune serum was also tested.

2.3.3.9 Western blot analyses

Twenty-four hours after transfection, cells were washed twice with PBS, lysed (Tris-HCl pH 7,4 50 mM, NaCl 120 mM, EDTA 5 mM, NP40 0,5%, Na₃PO₄ 0,2 mM, DTT 1 mM, NaF 50 mM) in the presence of protease inhibitors (Complete, Boehringer) and incubated on ice. Cell debris were pelleted by centrifugation. Protein concentration was determined by Bradford (BioRad). Samples were loaded onto 12% Bis-Tris (Invitrogen) gels, subjected to electrophoresis at 200V in MOPS buffer and transferred (Ibnot gel transfer, Invitrogen) onto a PVDF membrane. Membranes were blocked in a 5% PBS-milk solution and incubated with the appropriate primary antibody overnight: anti-APH2 (SPY557) 1:1000, anti-GFP 1:1000

(Clontech 632380), anti-Myc 1:1000 (Upstate clone 4A6). Membranes were then washed and incubated either with anti-rabbit or with anti-mouse horseradish peroxidase-conjugated secondary antibodies (Healthcare) and developed using the SuperSignal West Pico Chemiluminescent substrate Kit (Pierce). To control for the amount of protein loaded per well, membranes were stripped with the Re-blot Plus Kit (Chemicon International), and reprobed with anti β -tubulin antibody (sc9104 Santa Cruz Biotechnology 1:1000).

2.3.3.10 Immunofluorescence analyses

HeLa cells were analyzed by fluorescence microscopy at 24h post-transfection as previously described (Meertens et al., 2004a). Briefly, cells were washed with PBS, fixed, and permeabilized with 4% paraformaldehyde and 0.1% Triton X-100 for 30 min at room temperature. Cells were then incubated with a rabbit anti-nucleolin antibody (Affinity Bioreagents PA3-16875) for 1 h at room temperature. Samples were next washed with PBS and incubated with a FITC-conjugated mouse anti-Myc antibody (9E10, Sigma) and a secondary Dylight 549-conjugated anti-rabbit IgG (Pierce) antibody for 45 min at room temperature. Coverslips were mounted with Vectashield containing DAPI (Vector). Images were acquired with an Upright microscope Zeiss Axioplan 2 imaging, with the Zeiss Axiovision 4.4 software and with the Zeiss ApoTome system. This system provides an optical slice view reconstructed from fluorescent samples, using a series of "grid projection" (or "structured illumination") acquisitions.

2.3.3.11 Luciferase assays

293T cells were transiently transfected with pHTLV-2-LTR-luc or pHTLV-1-LTR-luc plasmids together with pSG5M-Tax2 or pSG5M-Tax1 (Meertens et al., 2004a) and APH-2-Myc or HBZ-SP1-Myc (Cavanagh et al., 2006) as previously described (Meertens et al., 2004b). For each transfection, the amount of transfected DNA was equalized by addition of a pSG5M backbone vector. All transfections were carried out in the presence of a phRG-TK vector (10ng) in order to normalize the results for transfection efficiencies. Reporter activities

were assayed 24 h post-transfection using the Dual-Luciferase reporter assay system (Promega). Luciferase activities were measured with a Berthold LB9500C luminometer as previously described (Meertens et al., 2004b).

2.3.4 Results

2.3.4.1 Existence of a putative ORF in the HTLV-2 minus strand.

In order to determine whether the HTLV-2 genome also contains a putative open reading frame on the minus strand, we first performed an analysis on HTLV-2 proviral DNA sequences available in GenBank, i.e. Gab, K96, AF412314, SP-WV, G2, G12, RP329, Efe2. This revealed the presence of a putative ORF in all of the analyzed strains except for MO, which had been used for generating the infectious molecular clone pH6neo (Shimotohno et al., 1985). We therefore re-sequenced this proviral DNA and discovered that the GenBank MO sequence (accession number M10060) contains a series of errors. After correction, it appeared that MO also possessed a similar antisense ORF. As HTLV-1 HBZ, this ORF is located between the Tax and Env ORFs in the HTLV-2 genome (Figure 2.1A). The putative protein encoded from this ORF was therefore named APH-2 for "Antisense Protein of HTLV-2"

2.3.4.2 Detection of a spliced APH-2 transcript in HTLV-2 transfected cells.

Given the similarities between the whole genomic organization of HTLV-1 and HTLV-2, we reasoned that an HTLV-2 antisense transcript encoding APH-2 would initiate from the 3' LTR and be spliced as it is the case for HTLV-1 HBZ SP-1. Antisense primers spanning different positions in the 3' LTR were thus designed and RT-PCR experiments were then conducted with RNA extracts obtained from 293T cells transfected with HTLV-2 proviral DNA. These experiments were performed either with the pH6neo vector or with pHTLV-2 Δ EcoRI, which is deleted from nt 1 to 4683 (Figure 2.1A). This latter construct can therefore not initiate sense transcription and is assumed to lead to increased antisense transcription. RT-PCR was first performed using different sets of primers on 293T cells transfected with pHTLV-2 Δ EcoRI. As seen in Figure 2.1B (lanes 1, 3), these RT-PCR analyses clearly demonstrated that antisense transcription is occurring in transfected 293T cells. Moreover, the size of the specific amplicons (Figure 2.1B and D) was shorter than

expected (564 bp instead of 1935 bp and 378 bp instead of 1749 bp, respectively). These results are indicative of a potentially spliced *APH-2* RNA and indicated that the unspliced transcript was in very low amount. With a more downstream forward primer, i.e. primer 21-1, no signal was detected for either a spliced or unspliced transcript (data not shown). RT-PCR analyses performed with primers 22-1/21-2 on RNA extracts obtained from cells transfected with the pH6neo clone confirmed the presence of a spliced *APH-2* mRNA (Figure 2.1C lane 2). Although this technique is not quantitative, and as previously reported for HTLV-1 HBZ, the RT-PCR signal was stronger in RNA extracts obtained from pHTLV-2ΔEcoRI transfected cells (Figure 2.1C lane 2 vs. 4). In all these RT-PCR experiments, no signals were detected when RT was omitted from the reaction mix (Figure 2.1B lanes 2, 4 and 2.1C, lane 1 and 3).

PCR products obtained in Figure 2.1B lanes 1 and 3 were then sequenced. Sequence analyses confirmed that they correspond to antisense transcripts that initiated in the 3' LTR and were spliced to the 5' end of the deduced *APH-2* ORF (Figure 2.1D). Splice donor (SD) and splice acceptor (SA) sites were located at position 8544 and 7173 of the antisense strand of pH6neo proviral DNA, respectively. The full length ORF region of *APH-2* would therefore span 1923 bp of the HTLV-2 minus strand with an intronic region of 1371 bp. Interestingly, analyzing the full length genome sequences of various HTLV-2 and STL-2 strains present in the GenBank database revealed that the SD and SA consensus sequences (underlined in Figure 2.2A) were always conserved, with the exception of the G2 strain which showed a slight variation in the SD consensus sequence. The ATG initiation codon positioned in exon 1 (in bold in Figure 2.2A) was also conserved in the different tested strains including those of STL-2. As depicted in this figure, the resulting spliced transcript should lead to a protein with 4 amino acids derived from exon 1 added to the remaining *APH-2* amino acid sequence. The amino acids bordering the splice junctions were also well conserved although occasional variations in amino acids were noted for the HTLV-2 and STL-2 strains. The expression of an unspliced transcripts would potentially lead to the expression of a protein with a different N-terminal structure (Figure 2.2B).

Altogether, these results demonstrate that antisense transcription is taking place in cells transfected with the HTLV-2 molecular clone. It initiates in the 3'LTR and produces a spliced transcript, which could be translated into a protein. As the spliced transcript seemed

to be the most important form of *APH-2* mRNA (versus an unspliced form, which was not detected in our analyses), the rest of this study was thus focused on this *APH-2* transcript and its encoded protein.

2.3.4.3 *APH-2* mRNA is present in HTLV-2 infected cells and in HTLV-2 carriers in vivo.

We next tested whether the *APH-2* spliced transcript was also present in two HTLV-2-infected cell lines (MO and C19) or in non-cultured PBMCs obtained from 5 different HTLV-2 blood donors. A semi-nested RT-PCR was thus performed on RNA from these various cell samples using primers located in exons 1 and 2, leading to an expected signal of 378 bp. *APH-2* mRNA was indeed detected in both HTLV-2-infected cell lines (Figure 2.3A, lanes 2 and 3) and in one out of five tested HTLV-2 carriers (Figure 2.3B, lane 9). No signal was amplified in RNA extracts obtained either from HTLV-1-infected cell lines (Figure 2.3A, lane 6) or from uninfected Jurkat and CEM cells (Figure 2.3A, lanes 4-5). Of note, due to the limited amount of RNA extracted from the PBMCs of the HTLV-2 carriers following DNase I treatment, RT-PCR was performed with a quantity of RNA ranging from 3 to 300 ng/sample. This might explain why *APH-2* was only detected in one sample. The RT-PCR products were cloned and sequenced and were demonstrated to be genuine *APH-2* cDNAs originating from spliced transcripts (data not shown).

2.3.4.4 HTLV-2 antisense transcription initiates at different positions in the 3'LTR.

We were then interested in determining the transcription initiation sites of the *APH-2* transcript. Following transfection of 293T with pHTLV-2ΔEcoRI, RNA was analyzed using the SMART RACE kit and a reverse primer positioned in exon 2, close to the splice junction. Cloning and sequencing of the amplified products generated by 5' RACE (Fig. 2.4A) led to the identification of several major initiation sites positioned in the 3' LTR (exclusively in the R and U5 regions) and spanning a total of 210 nt (Fig. 2.4B)). The position of these initiation sites is in agreement with the size of the signal observed in Figure 2.4A. These results hence

confirmed that *APH-2* transcripts initiate in the 3' LTR at multiple positions. This multiplicity of initiation sites might be a consequence of the absence of a TATA box at close proximity.

2.3.4.5 Position of the poly A site.

The *APH-2* mRNA 3' end was next mapped using 3'RLM-RACE that specifically amplifies the 3' end of polyadenylated RNA. The pHTLV-2ΔEcoRI plasmid was transfected into 293T cells and RNA was extracted 24h later. Using a primer positioned downstream of the *APH-2* stop codon, we amplified a 590 bp fragment (Figure 2.5A). Sequencing of this fragment demonstrated that the poly A tail was positioned at a distance of 1449 nt from the *APH-2* stop codon. As in the HTLV-1 HBZ transcript (Cavanagh et al., 2006), the cleavage site needed for the addition of the polyA tail was located in a TA dinucleotide positioned 17 nucleotides downstream of the polyA signal and a few nucleotides from a GT-rich segment, which represents another typical consensus sequence for poly A addition (Figure 2.5B). These consensus sequences were highly conserved among other various HTLV-2 strains (Figure 2.5B).

To confirm the above characterization of the *APH-2* transcript, we then performed a Northern blot analysis on extracts obtained from 293T cells transfected with the pHTLV-2ΔEcoRI plasmid. The coding region of the *APH-2* cDNA was used as a probe. Interestingly, a signal lower than 2.5 kb was specifically detected in HTLV-2-transfected cells but not in non-transfected cells (Figure 2.5C, lane 1 vs 2). The size of the transcript corresponds to the 2.3 kb predicted *APH-2* mRNA when taking in consideration both 5' and 3' RACE data.

Altogether, these results demonstrate that *APH-2* mRNA initiates at different positions in the 3'LTR, is spliced and polyadenylated. Analysis of the resulting *APH-2* protein sequence (Figure 2.6) demonstrated that it does not contain a bZIP domain unlike HBZ (Gaudray et al., 2002) but interestingly possesses a LXXLL sequence at its COOH-terminus (amino acids 179-183) as well as a central IXXLL (amino acids 64-68) sequence. These domains were recently shown to be involved in the binding of HTLV-1 HBZ to the KIX domain of CBP/p300 and to interfere with the ability of HTLV-1 Tax to bind to these transcription co-activators (Clerc et al., 2008).

2.3.4.6 APH-2 and HBZ-SP1 do not have a similar localization in transfected cells.

APH-2 cDNA was cloned into the pcDNA3.1-Myc expression vector, allowing the production of a Myc-tagged APH-2 protein (referred as APH-2-Myc throughout the text). We then transfected APH-2-Myc in HeLa cells. In these experiments, a HBZ-SP1-Myc expression vector was used as a control. Detection of the tagged proteins was achieved by incubation of transfected cells with an anti-Myc antibody (Figure 2.7A). A strong nuclear signal was observed in cells transfected with the APH-2-Myc construct (top panel). However, the APH-2-Myc protein was not nucleolar as shown with the anti-nucleolin staining. This was confirmed by the Apotome technique (Figure 2.7B top panel). Unlike APH-2, and as previously published, HBZ-SP1 was nucleolar (Figure 2.7A lower panel and Figure 2.7B bottom panel). Altogether, these results demonstrate that both viral proteins are nuclear but that only HBZ-SP1 localizes to the nucleolus. Similar results were obtained with GFP-tagged APH-2 and HBZ-SP1 constructs in HeLa and COS cells (data not shown).

Detection of the APH-2-Myc protein was also performed by Western blot analysis either with anti-APH-2 or anti-Myc antibodies (Figure 2.7C). Both antibodies led to the detection of an APH-2 protein having an approximate 25 kDa size. As a control, HBZ-SP1-Myc was detected with the Myc antibody, but not with the anti-APH-2 antibody. The polyclonal APH-2 antibody is therefore specific of APH-2 and can be used in Western blot experiments.

2.3.4.7 APH-2 down-regulates Tax2-mediated HTLV-2 LTR activity.

Previous reports demonstrated that HBZ-SP1 down regulates HTLV-1 transcription by interacting with CREB-2 and CREB through its bZIP domain. HBZ-SP1 also interacts with CBP/p300 co-activators through LXXLL-like motifs and prevents their recruitment on the viral promoter. Therefore, at least two domains of HBZ-SP1 play a role in the down-regulation of viral expression. As stated above, APH-2 does not contain a bZIP motif and is therefore unlikely to bind CREB or/and CREB-2. However, APH-2 contains a typical LXXLL sequence. We therefore hypothesized that, despite the lack of a bZIP domain, APH-2

could act as a viral repressor. 293T cells were transfected with an HTLV-2 LTR reporter construct, constant amount of the HTLV-2 Tax protein and increasing amount of the APH-2 expression vector (Figure 2.8A top panel). We could reproducibly observe that APH-2 exerted a negative effect on Tax2-mediated transcription and that this effect was dose-dependent, while APH-2 had no effect in the absence of Tax2. As a control, a similar experiment was conducted with HBZ-SP1 and showed a similar repression on Tax1-induced HTLV-LTR activity (Figure 2.8A bottom panel). The experiments were repeated in Jurkat T-cells and gave similar results (data not shown). The levels of expression of Tax1, Tax2, HBZ-SP1 and APH-2 were determined as a control (Figure 2.8B).

2.3.5 Discussion

It has been accepted for a long period of time that, in retroviruses, transcription only initiates in the 5'LTR. The existence of antisense transcription was therefore controversial (Larocca et al., 1989; Vanhee-Brossollet et al., 1995). However, the demonstration that retroviral antisense transcription exists has recently been strongly supported for two human retroviruses i.e. HTLV-1 and HIV-1 (Gaudray et al., 2002; Landry et al., 2007). In HTLV-1, the HBZ protein, which is encoded from such an antisense transcript, is currently the subject of numerous studies. This protein binds several transcription factors (CREB, CREB-2, c-Jun, JunB, JunD) or transcription co-activators (CBP/p300) and alters their functions (Basbous et al., 2003; Clerc et al., 2008; Hivin et al., 2006; Hivin et al., 2007; Lemasson et al., 2007; Matsumoto et al., 2005; Thebault et al., 2004). Interestingly however, HBZ transgenic mice in which the transgene is expressed under the CD4 promoter/enhancer do not develop severe pathologies but demonstrate an increased percentage of CD4⁺ T lymphocytes in the splenocytes and display a higher proliferation of this cellular population (Satou et al., 2006).

Given the fact that HTLV-2-infected individuals have a persistent lymphocytosis (Ed Murphy, unpublished observations), we wondered whether HTLV-2 could also encode a protein that would be functionally related to HBZ. The existence of an "HBZ-like" protein in HTLV-2 has been ruled out by a number of authors (Gaudray et al., 2002; Li and Green, 2007), while others suggested that such a protein might exist in HTLV-2 (Switzer et al., 2006). These discrepancies are due to two facts: (i) the HTLV-2 prototype strain (MO) (Shimotohno et al., 1985) that has been used for searching an antisense ORF contains several sequence errors that prematurely interrupt the putative ORF and (ii) there is an antisense ORF in the other HTLV-2 strains but its sequence does not contain a bZIP domain. Therefore, on a semantic point of view, it is correct to say that an HBZ-like antisense sequence, is absent from the HTLV-2 genome.

In this report, we first showed that a spliced mRNA corresponding to APH-2 is expressed *in vitro* in cells transfected with the pH6neo molecular clone and in all HTLV-2-infected cell lines tested (regardless of the viral subtype). This transcript is structured similarly to the HBZ transcript i.e. it is spliced, initiates in the 3' LTR at multiple sites, and is polyadenylated (Figure 2.9). The length of the intron and of the 5' and 3' untranslated regions

is also similar to that of the HBZ transcript. These similarities might argue that the expression of antisense transcription in the HTLV (and STLV) retrovirus family has been conserved among the different viruses. The positioning of both APH-2 and HBZ ORFs which is similar might be favored to avoid undesirable antisense effect on Tax2 sense transcripts. The fact that, unlike *HBZ* mRNA, we did not observe alternative splicing for HTLV-2 *APH-2* mRNA remains to be confirmed. Our results also suggest that unspliced *APH-2* transcript is very low in abundance and could not actually be detected in our RT-PCR analysis. This is in contrast to HBZ, which produces a substantial amount of unspliced transcript potentially encoding for a different HBZ isoform (Murata et al., 2006; Usui et al., 2008; Yoshida et al., 2008). However, further studies are required to indeed confirm that unspliced *APH-2* transcripts is poorly detectable in HTLV-2-infected cells as such a transcript could indeed encode for a different APH-2 isoform.

A recent report clearly demonstrated that HTLV-1 *HBZ-SP1* is expressed *in vivo* in almost all HTLV-1-infected individuals (healthy carriers and ATLL patients) and that its level was correlated with the proviral load (Usui et al., 2008). Our current results, although obtained from a limited series of HTLV-2 carriers indicate that *APH-2* is expressed *in vivo*. The lack of detection in some carriers is likely due to the very limited amount of RNA that was available for the RT-PCR experiments. A study based on a much larger number of HTLV-2 individuals using qPCR and q-RT-PCR techniques is currently underway to address this issue. This study will allow us to examine whether the HTLV-2 proviral load, the levels of *APH-2* and *Tax2* transcripts are correlated. Whether APH-2 is expressed at the protein level in these cell samples from infected patients and elicits a humoral or a cellular response is not known and will also be investigated.

Our results have led us to demonstrate that differences between HBZ SP1 and APH-2 were apparent in terms of their nuclear localization. Indeed, HBZ-SP1 is specifically retained in nucleolar structures, unlike APH-2. Differences at the levels of protein-protein or nucleic acid-protein interactions might contribute to these varying nuclear localization and could be dependent on both differences in amino acid sequence and/or in post-translational modifications. Additionally, a Nucleolar Localization Signal might be specifically present in HBZ-SP1. Future studies will be required to discriminate between these possibilities.

Despite its lack of a bZIP consensus sequence, APH-2 prevents transcription from the HTLV-2 5'LTR through an unknown mechanism. It would be therefore of interest to find out whether APH-2 is able to recruit and to squelch cellular proteins that are required for viral transcription. As stated above, a recent report clearly established that the two LXXLL-like motifs that are located within the *N*-terminal sequence of HBZ-SP1 allow the protein to bind the CBP/p300 co-activators and to squelch them (Clerc et al., 2008). Interestingly APH-2 possesses one LXXLL motif at its COOH-terminus as well as a central IXXLL sequence. Whether one or both of these motifs are responsible for the APH-2 inhibitory function on Tax2-mediated LTR activation remains to be determined. Alternatively, we cannot exclude the fact that *APH-2* mRNA might also play a role, as suggested for HBZ-SP1 (Satou et al., 2006; Yoshida et al., 2008).

In conclusion, we have demonstrated that the HTLV-2 minus strand is also transcribed and that the encoded APH-2 protein shares some functional similarities with HTLV-1 HBZ even if, unlike HBZ, APH-2 does not contain a bZIP domain. Interestingly, the latter has been found to bind to the different Jun transcription factor family members that are involved in the control of different target genes. These genes are important in regulating a series of biological processes including proliferation, differentiation, apoptosis, and transformation. Nevertheless, it is still possible that despite the lack of this sequence, APH-2 is accountable for the cell proliferation which leads to the persistent lymphocytosis observed among HTLV-2 carriers. If this is the case, the lack of hematological disease could be solely due to the intrinsic properties of Tax2.

2.3.6 Acknowledgments

RM is supported by INSERM. ED and CJ are supported by the Ministère de la Recherche. NLK is supported by the Croucher Foundation. This work was supported by a grant to J.M.M. from the Agence Nationale de Recherches sur le Sida et les hépatites virales (ANRS) and a grant to B.B. from the Cancer Research Society (CRS) Inc. I.C. is supported by ANRS. M.H. is supported by an institutional Hydro-Quebec scholarship and S.L. holds a CIHR PhD scholarship. B.B. holds a Canada Research Chair of Canada (Tier 2). We thank Mrs Emmanuelle Perret (www.pfid.org) from the Plateforme d'Imagerie Dynamique, Imagopole, Institut Pasteur for her help with the imaging.

2.3.7 Figure legends

Figure 2.1: HTLV-2 antisense transcripts are detected and are spliced.

(A) The HTLV-2 molecular clone pH6neo provirus is depicted with all known genes and is adapted from (Feuer and Green, 2005). The positioning of the putative APH-2 ORF from the antisense strand is indicated below the proviral DNA. A 5' deleted version of this vector termed pHTLV-2 ΔEcoRI is also presented and encompasses nt 4683 to 8952. The primers used for the RT-PCR experiments are also represented with their positioning. The dashed line represents the hypothetical antisense transcript. (B,C) 293T cells were transfected either with pHTLV-2 ΔEcoRI (B,C) and/or with pH6neo (C). RNA was extracted and analyzed by RT-PCR for the presence of antisense transcripts using the primer combinations 19-1/21-2 and 22-1/21-2. (C): the 22-1/21-2 primer set that allows the amplification of the spliced transcript was used. When indicated, RT was omitted during the RT step. Control= no cDNA; M= 100 bp marker. (D) The position of exon 1 and exon 2 of APH-2 transcripts (black boxes) is presented. The size of the RT-PCR products are also indicated.

Figure 2.2: The splice donor and acceptor sequences of the APH-2 transcript are conserved.

(A) DNA sequences containing the APH-2 splice donor and splice acceptor sites were compared between various HTLV-2 and STLV-2 isolates retrieved from GenBank. The correct sequence of the MO proviral DNA was used as a reference. Bold letters: exonic sequence. Underlined: splice donor (SD) and splice acceptor (SA) sequences. The HTLV-2 subtype is indicated on the left of the panel. Predicted amino acid sequence of the spliced *APH-2* transcript at the splice junction are also presented above the exonic nucleotidic sequence (B). Predicted amino acid sequence of the spliced *APH-2* transcript at the splice junction and comparison with the amino acid sequence of the unspliced transcript.

Figure 2.3: The spliced *APH-2* mRNA is detected in HTLV-2-infected cell lines and in PBMCs from infected blood donors.

(A) Total RNA was extracted from C19 and MO (HTLV-2 infected), Jurkat and CEM (HTLV non-infected) and C91/PL (HTLV-1 infected) cell lines. RT was performed using the 24-2 primer and was followed by a semi-nested PCR using the 19-1/21-2 and 22-1/21-2 primers. Lanes 1-2: C19; Lane 3: MO; Lane 4: Jurkat; Lane 5: CEM; Lane 6: C91/PL; Lane 7: H₂O. (B) Same experimental conditions as in (A). Lanes 2 and 3: 400 ng and 40 ng of RNA extracted from MO cells were used for RT-PCR analysis respectively. Lane 4: H₂O; Lanes 5-9: RNA extracted from five HTLV-2 infected individuals was used for RT-PCR. As a control (bottom), the same samples were analyzed for the presence of *GAPDH* mRNA by RT-PCR. (A; B): lane 1: RT was omitted during the RT step. M= 100 bp marker.

Figure 2.4: HTLV-2 antisense transcription initiates in the 3' LTR.

(A) 5'RACE PCR analysis was conducted using RNA samples from 293T cells transfected with pHTLV-2ΔEcoRI (lane 2). The cDNA was synthesized with a modified oligodT and ligated to a supplied anchor at their 5' end. Two subsequent PCR rounds were conducted with the 5'RACE outer and inner primers along with the transcript-specific 24-2 and 21-2 reverse primers. As a negative control, final PCR amplifications were performed with water (no cDNA) (lane 1). The resulting amplified products were run on an agarose gel. M = 100 bp marker. (B) The position of the identified transcription initiation sites of *APH-2* mRNA is depicted in the 3' LTR. Nucleotide positioning is relative to the *APH-2* ATG initiation codon numbered as +1. The numbering in the proviral DNA correspond to the sense strand.

Figure 2.5: Identification of the poly A addition site of the *APH-2* transcript.

(A) RNA samples obtained from 293T cells transfected with the pHTLV-2ΔEcoRI plasmid (lane 2) were analyzed by 3' RLM-RACE using the 21-3 primer. Lane 1: PCR amplification in the absence of any cDNA. M: 100 bp marker. (B) Position of the poly A addition site next to a consensus poly A signal and a GT-rich consensus sequence. The sequences of the *APH-2*

mRNA and of the 3' poly A tail are shown. The black boxes represent the coding portion of the APH-2 spliced transcript. HTLV-2 sequences retrieved from GenBank were also compared using MO as a reference sequence. Comparisons were focused on the AATAAA polyA signal (position 5185 on the sense strand), the cleavage site deduced from our 3'RACE results and the GT-rich sequence (underlined in the pH6neo proviral DNA sequence). (C) PolyA⁺ RNA was extracted from pHTLV-2ΔEcoRI-transfected 293T cells (lane 1) or from non-transfected cells (lane 2) and was run on a 1.5% agarose gel. Following migration and transfer, the membrane was hybridized with a probe corresponding to the coding segment of the APH-2 cDNA and signals were revealed with a PhosphorImager. A RNA marker was migrated in parallel and corresponding molecular weights are indicated on the left side of the panel.

Figure 2.6. Amino acid sequence of the APH-2 protein.

The amino acid sequence of APH-2 was deduced from the MO strain. Underlined are the peptides used for the immunization of the rabbits. Dashed boxes highlight two potential LXXLL motifs.

Figure 2.7: APH-2 localizes to the nucleus.

(A) APH-2-Myc (a-d) or HBZ-SP1-Myc (e-h) expression vectors were transfected into HeLa cells. Twenty four hours post-transfection, cells were fixed and stained as described in Material and Methods section. Cells were finally mounted in the presence of DAPI containing mounting medium. Images are representative of the entire population of transfected cells. Panels d and h: images are presented after merging of all three channels. (B) Same as in (A), panels d,h with the Apotome technique. (C). Cell lysates (50 μg) from 293T cells transfected as in (A) were subjected to electrophoresis on a 12 % Bis-Tris gel and analyzed by Western blot with anti-Myc, anti-APH-2, or anti-β-tubulin antibodies. pcDNA3.1/Myc parental vector was used as a negative control.

Figure 2.8: APH-2 represses Tax2 dependent transcription from the HTLV-2 LTR.

293T cells were transiently transfected with 125 ng of pHTLV-2-LTR-luc (top) or pHTLV-1-LTR-luc (bottom), 500 ng of pSG5M-Tax2 (top) or pSG5M-Tax1 (bottom) together with increasing amounts (125; 250; 500 and 1000 ng) of APH-2-Myc (top) or HBZ-SP1-Myc (bottom). The presented luciferase activities are an average of 3 (top) or 2 (bottom) independent experiments. (B) Fifty μ g of protein extracts from lysates obtained from transfected 293T cells were subjected to electrophoresis and probed for Myc, Tax2 or Tax1.

Figure 2.9. Comparison of HBZ and APH-2 transcripts.

HTLV-1 and HTLV-2 proviral DNA are presented along with most of their genes (accessory genes were omitted to simplify the figure). Previously reported unspliced and spliced HBZ transcripts are presented below the HTLV-1 genome. The APH-2 antisense transcript is also depicted next to the HTLV-2 genome. For each antisense transcript, dark areas highlight the coding regions of the transcript.

2.3.8 Figures

A

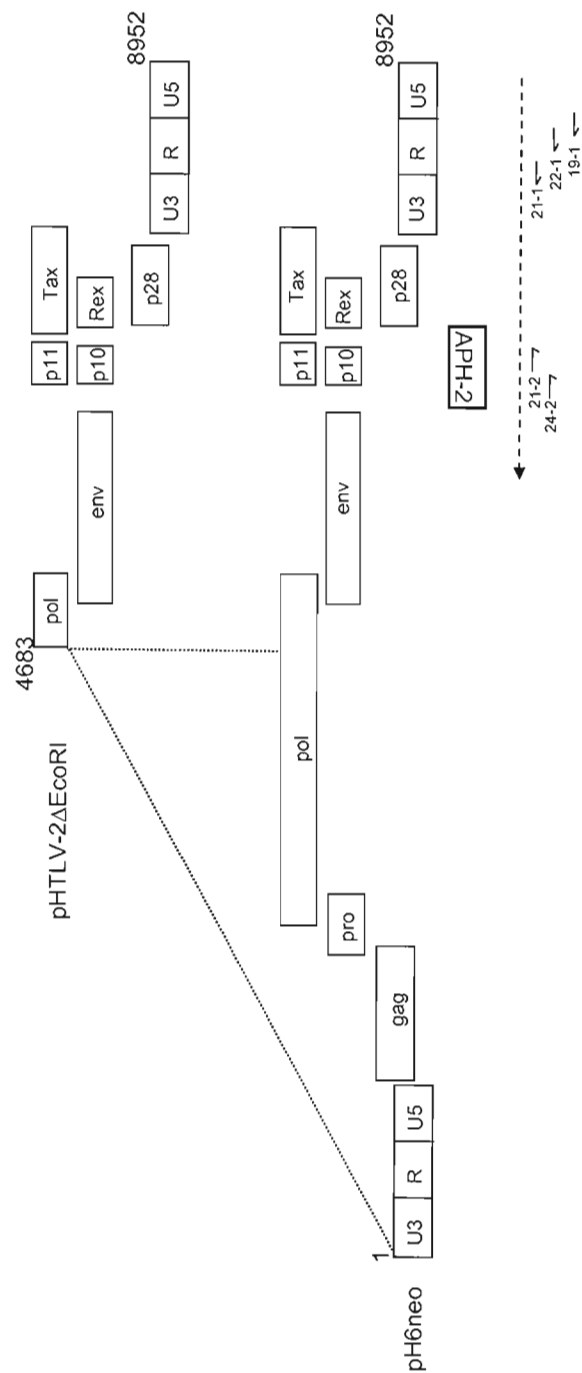
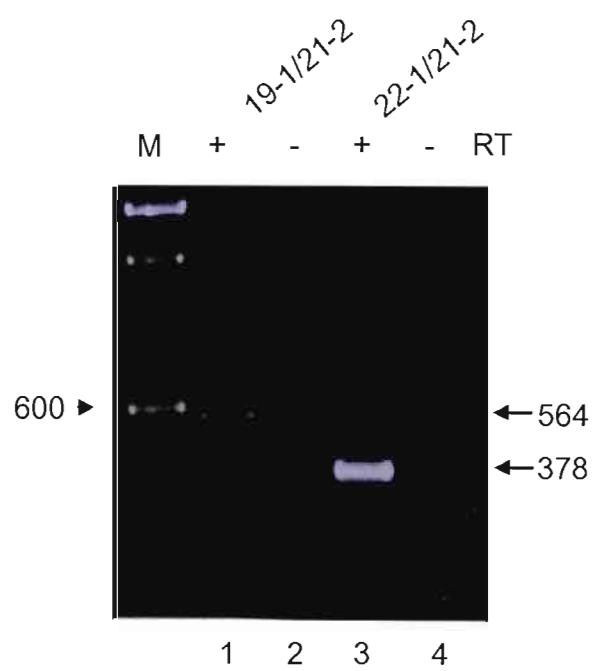


Figure 2.1

B



C

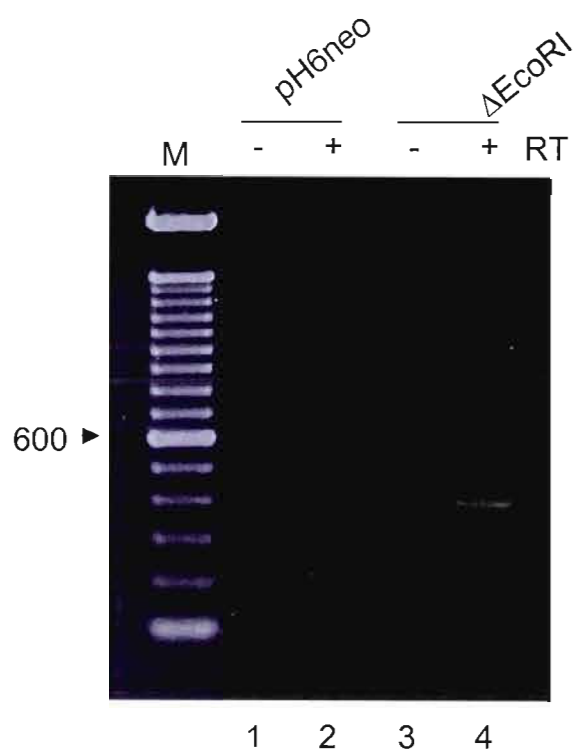


Figure 2.1

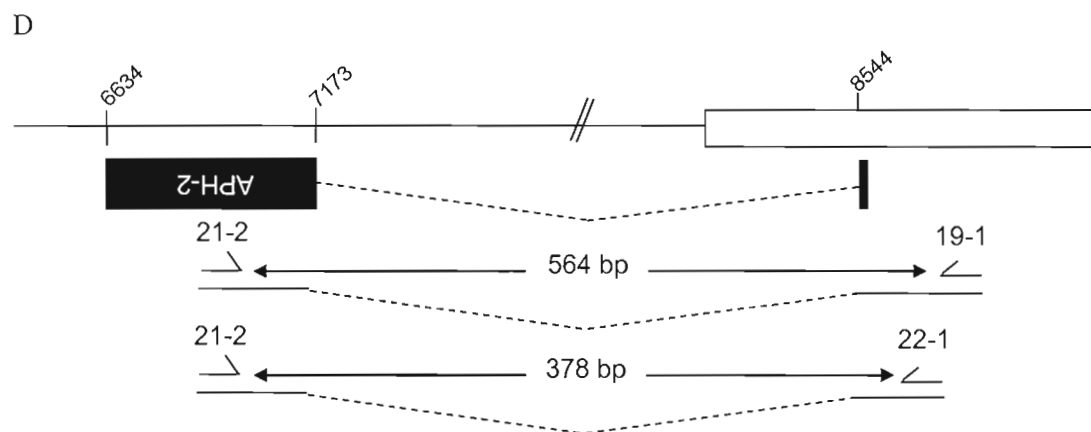


Figure 2.1

Figure 2.2

		Strain		Exon 1	Exon 2
HTLV-2 subtype				M V P Q	D Y F R
				ATGGATCCCCAAGGTGAGTCTCCGT	TGTGATTTTTTAGACTATTTTAGGA
				SD	SA
			A.....
			
			
			
			C.T.....C.	...A.....
			A.....
			
			C-A.....C
			T.....-..G.A.AGC	TATTTTG.....C....
			T.....-..G.A.AGC	TATTTTG.....C....

A

B

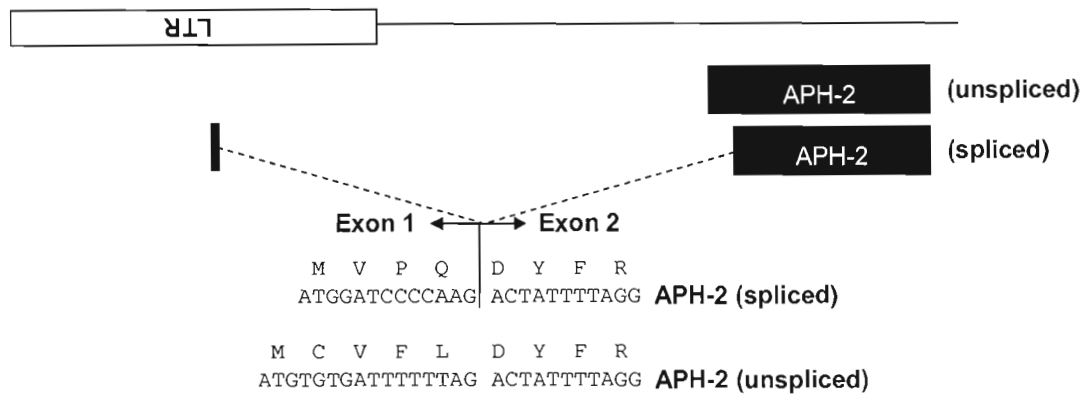
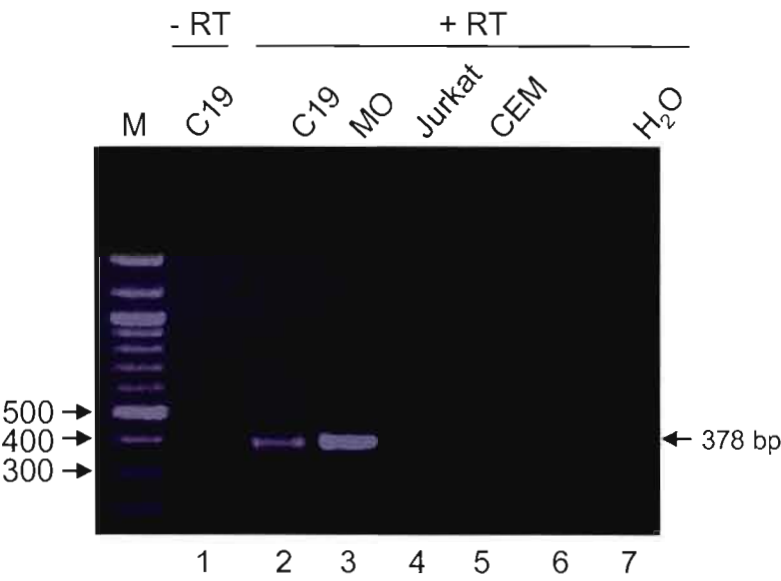


Figure 2.2

A



B

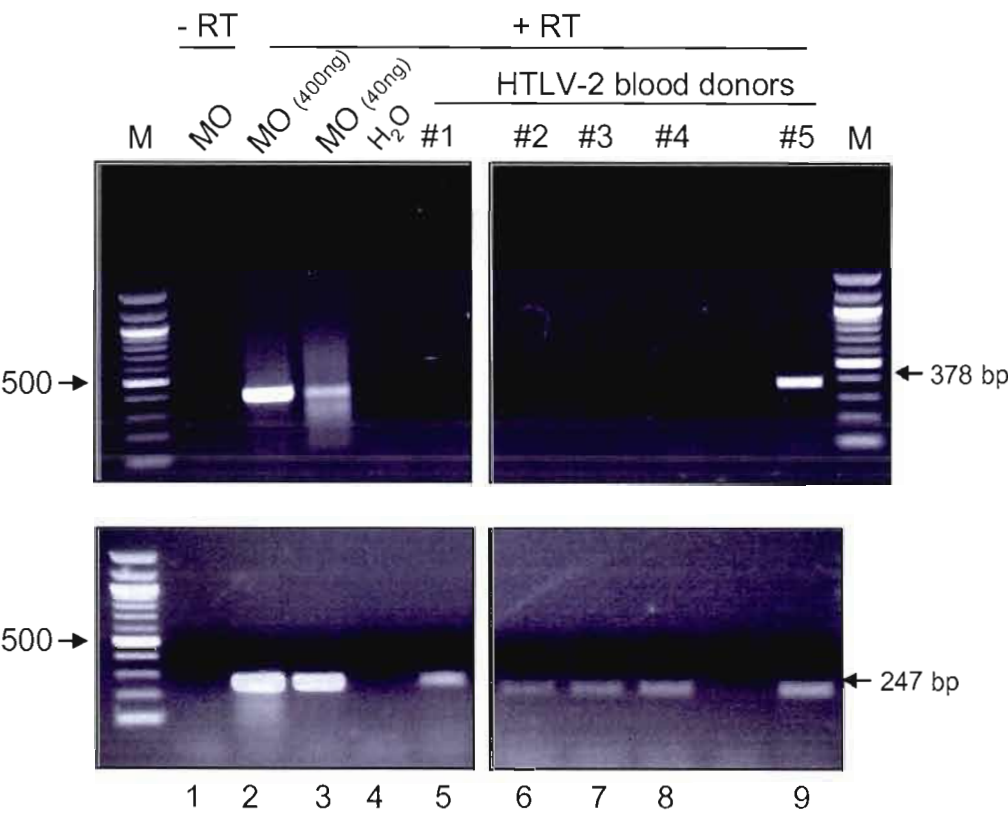
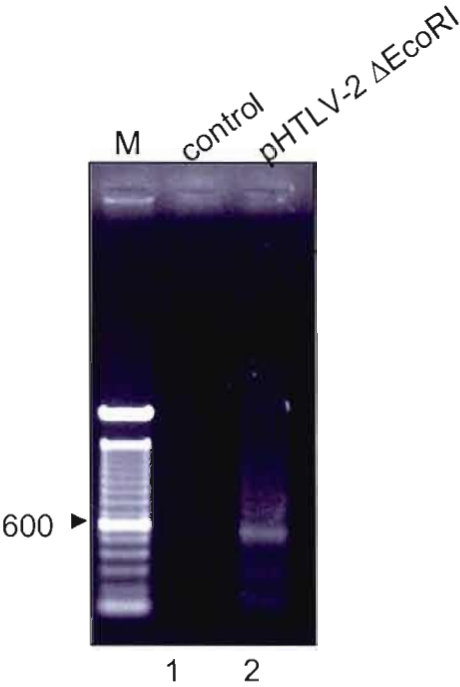


Figure 2.3

A



B

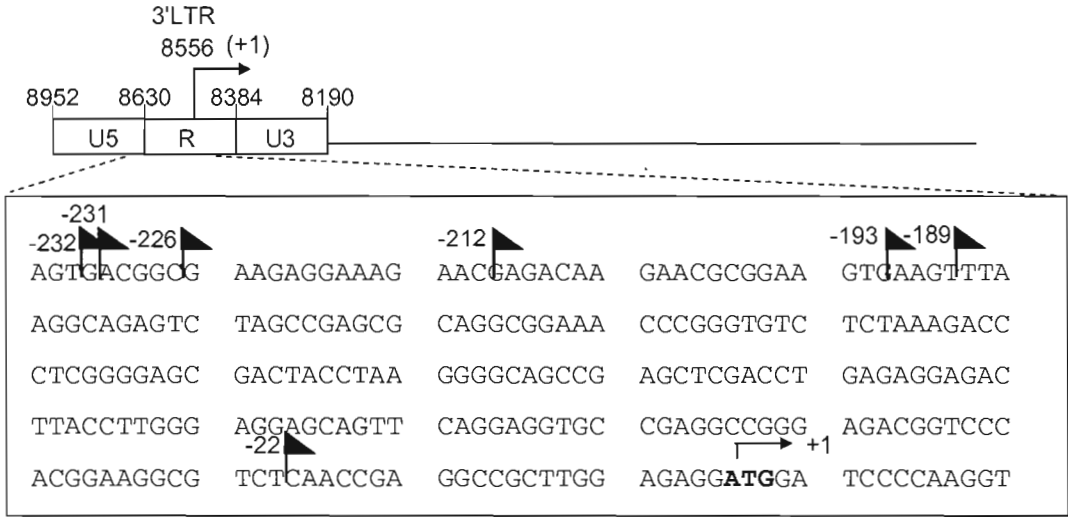
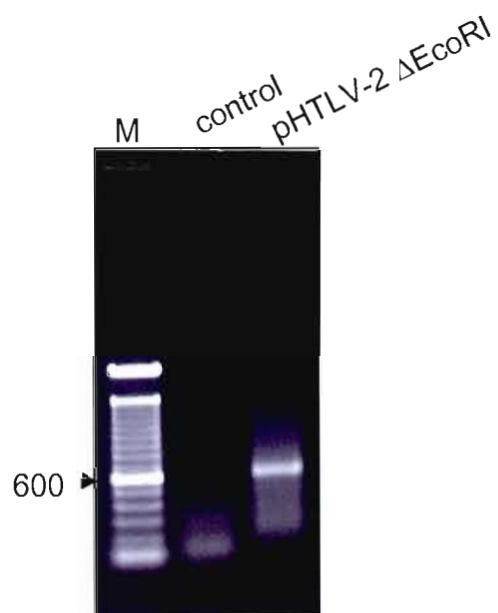


Figure 2.4

A



B

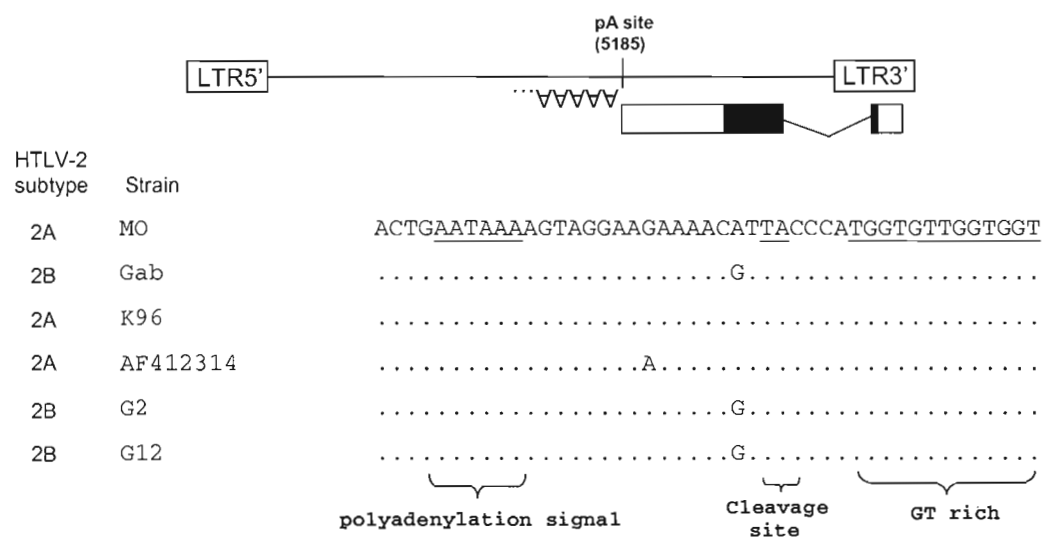


Figure 2.5

C

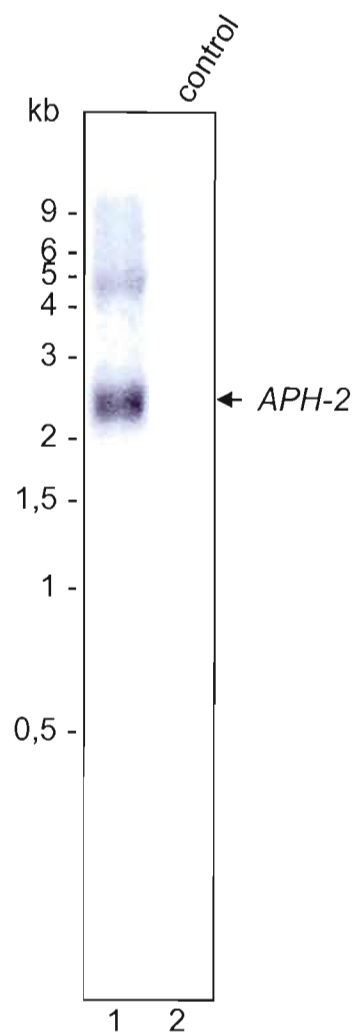


Figure 2.5

APH-2 sequence

₁MDPQDYFRRL PLGDQGEASL TSGVPPPGEE PRRTAKGRPL GSGSVSEQVR DSRRLLAEE QREILKLIQE REERERRRKE
RLREKDQKRR EREERHQQL E CIGLLGFDGF CDLLQGYVDF LEGEKKVLWE ECERGLEELF EAI IQFESGL DIGGDSSGIE
DCGWVNGRGT NHGDLLQKLA GLL₈₃

Figure 2.6

A

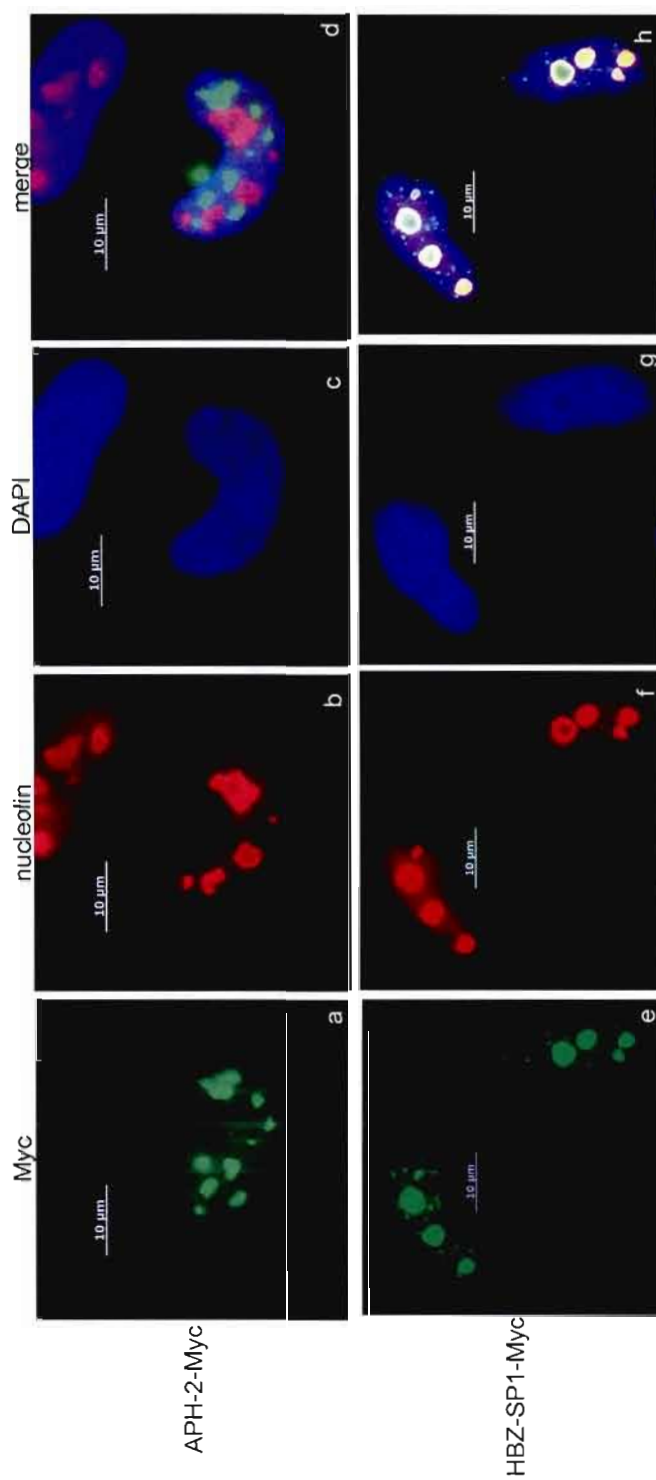
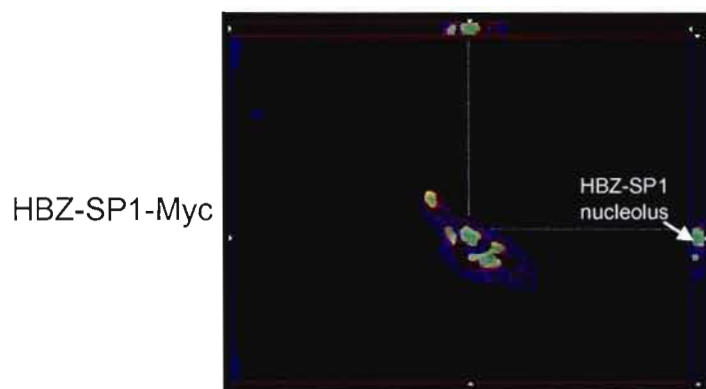
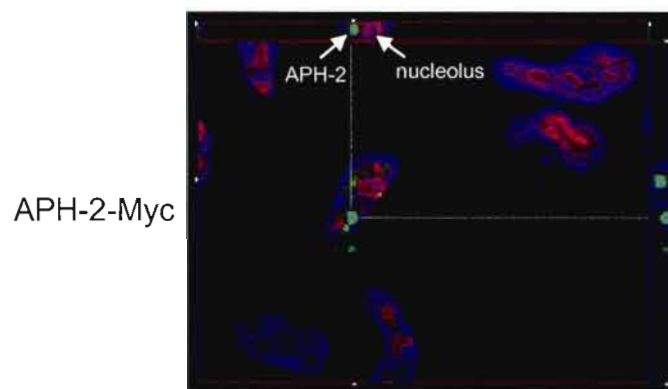


Figure 2.7

B



C

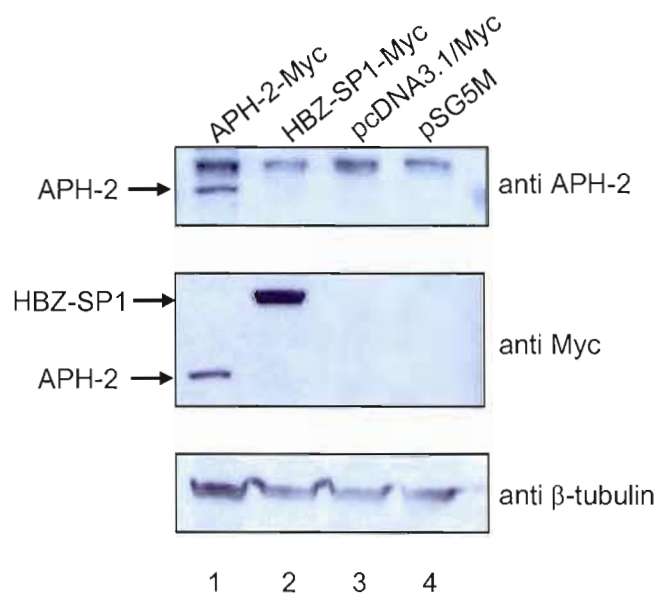


Figure 2.7

A

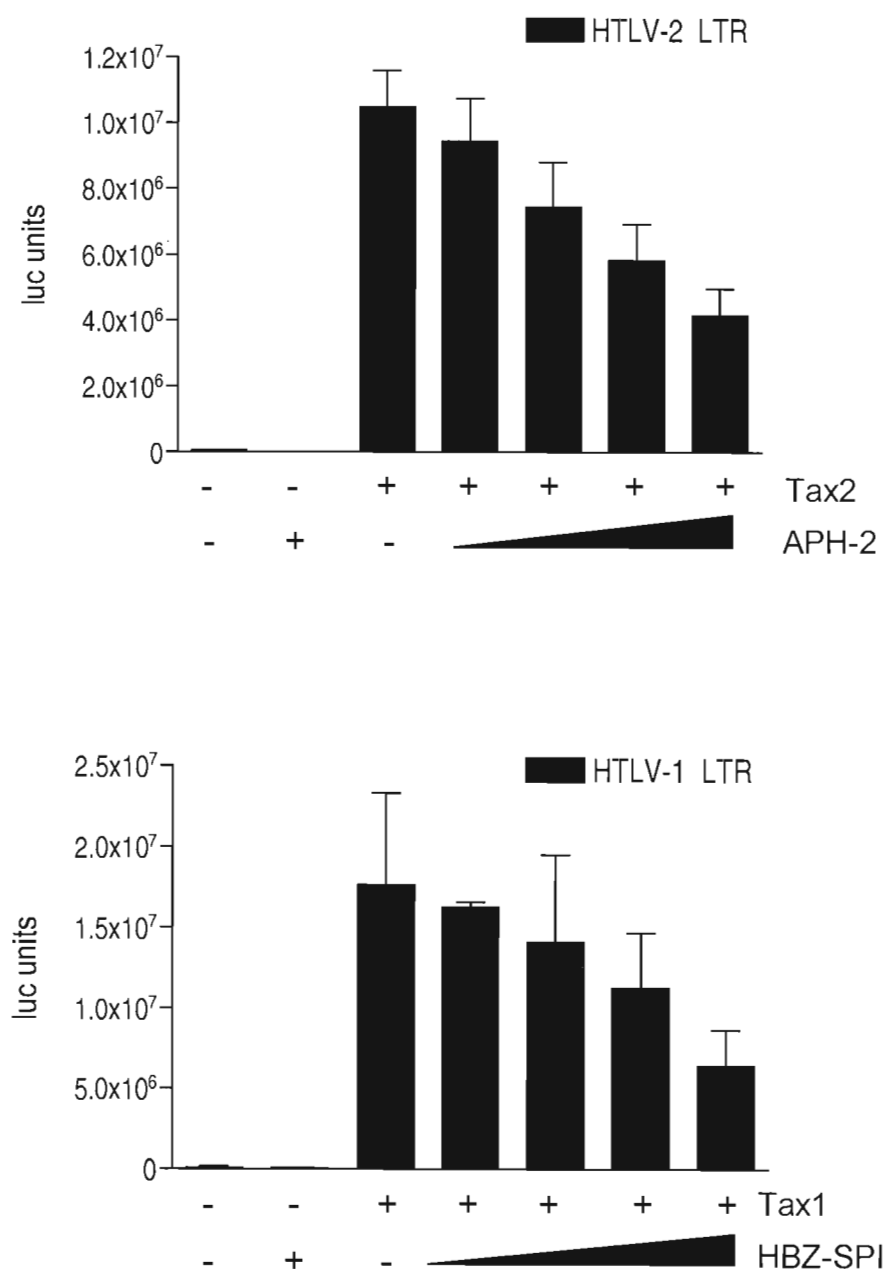


Figure 2.8

B

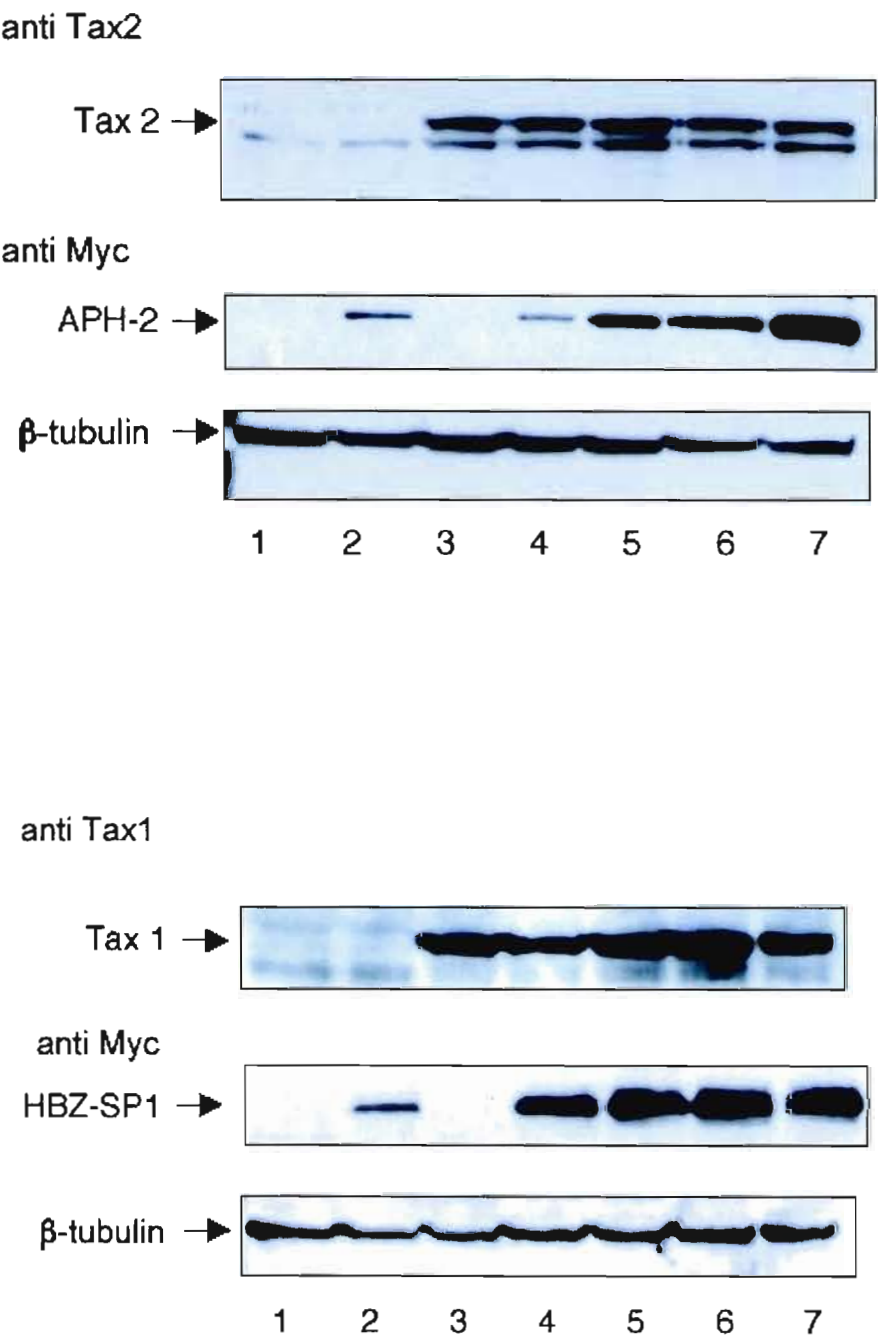
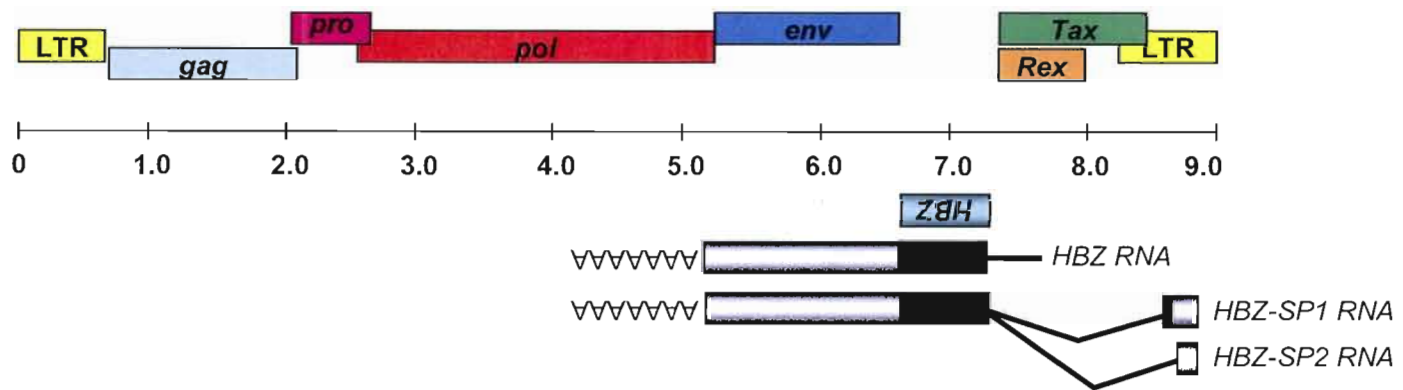


Figure 2.8

HTLV-1



HTLV-2

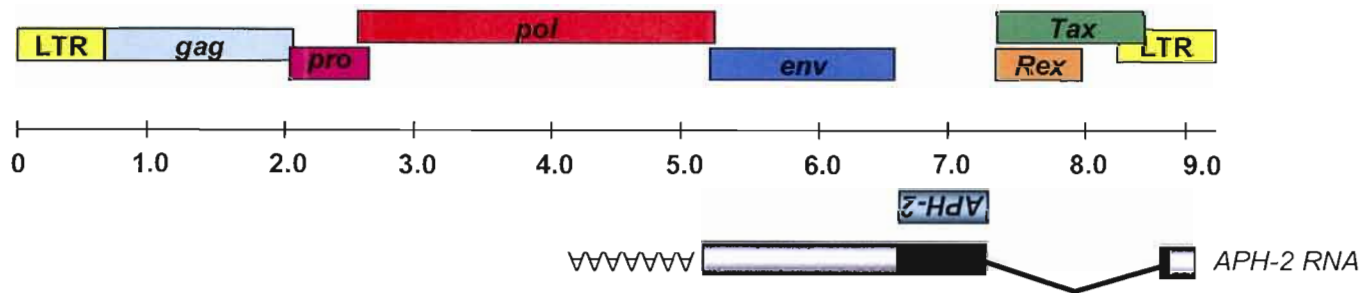


Figure 2.9

CHAPITRE III

ARTICLE 2

3.1 Contribution des auteurs

L'article a été rédigé par l'auteure du présent mémoire, Marilène Halin, avec les précieux conseils de son directeur de recherche, Benoît Barbeau. Marilène Halin a également effectué la majorité des expériences de RT-PCR, de RACE, de microscopie confocale, de Western blot ainsi que plusieurs transfections en plus de participer au montage des figures. Sébastien Landry a contribué aux expériences de RT-PCR, de microscopie confocale et il a effectué plusieurs transfections. Émilie Larocque a effectué des Western blot et elle a en outre aidé lors des expériences de microscopie. Suzan Marriott a fourni les vecteurs d'expressions de Tax 3 et Tax 4. William Switzer a fourni les constructions provirales complètes de HTLV-3 et HTLV-4. Benoit Barbeau a élaboré et mis sur pied le projet. Il a participé à l'écriture de l'article et au montage des figures.

La mise en forme de l'article correspond aux exigences de présentation des mémoires de l'Université du Québec à Montréal. L'article a été soumis au journal *Journal of virology*.

3.2 Résumé

Les rétrovirus humains HTLV-3 et HTLV-4 ont récemment été découverts et font partie de la famille HTLV, tout comme les rétrovirus HTLV-1 et HTLV-2. La transcription antisens chez les rétrovirus humains est un concept émergent qui a été clairement démontré, pour la première fois, chez le rétrovirus HTLV-1. En effet, le transcrit antisens d'HTLV-1 mène à la synthèse d'une nouvelle protéine nommée HBZ. Puisque des cadres de lectures ouverts sont conservés sur les brins complémentaires des génomes des rétrovirus HTLV-3 et HTLV-4, nous nous sommes intéressés aux potentiels codants de ces derniers. En effet, notre étude fait état de la détection d'une transcription antisens par des expériences de RT-PCR à partir de l'ARN de cellules 293T transfectées par l'ADN proviral complet des rétrovirus HTLV-3 ou HTLV-4. Ces transcrits sont épissés et polyadénylés. Les protéines produites par ces transcrits ont été nommées APH (*Antisense protein of HTLV*)-3 ou APH-4 et toutes les deux ne contiennent pas de domaine *leucine zipper* conservés; elles sont néanmoins constituées de régions basiques. Des expériences de microscopie confocale et d'immunobuvardage de type Western blot ont démontré qu'APH-4 est localisé uniquement au noyau, tandis que APH-3 semble à la fois nucléaire et cytoplasmique. L'étude de l'activité promotrice antisens a révélé que l'expression de APH-3 et APH-4 est modulée par des agents activateurs des cellules T. Finalement, ces deux protéines ont inhibé l'activation de la transcription virale de HTLV-1 médiée par les protéines virales Tax1 et Tax3. Ces résultats confirment que la transcription antisens n'a pas exclusivement lieu chez HTLV-1 et semble s'inclure dans la stratégie répliquative des rétrovirus HTLVs et que les protéines APH-3 et APH-4 jouent possiblement des rôles dans la réplication virale.

3.3 HTLV-3 and HTLV-4 viruses produce antisense transcripts encoding proteins with a Tax-inhibiting function

Marilène Halin¹, Sébastien Landry¹, Émilie Larocque¹, Suzan Marriott², William Switzer³ and Benoit Barbeau^{1*}

¹Département des Sciences Biologiques et Centre de recherche BioMed, Université du Québec à Montréal, 2080 Saint-Urbain, Montréal, Québec, Canada H2X 3X8, ²Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston TX 77030; ³Laboratory Branch, Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB prevention, Centers for Disease Control and Prevention, Atlanta, GA 30333.

*Correspondence to Benoit Barbeau, Département des Sciences Biologiques, Université du Québec à Montréal, 2080 Saint-Urbain Montréal, Québec, Canada H2X 3X8. Tel: 514 987-3000 ext.4576, E-mail: barbeau.benoit@uqam.ca

3.3.1 Abstract

Human HTLV-3 and HTLV-4 retroviruses have recently been discovered and are part of the HTLV family composed of the well-known HTLV-1 and HTLV-2 viruses. Antisense transcription in human retroviruses is a concept, which has recently emerged and has been conclusively demonstrated in HTLV-1, resulting in the synthesis of a viral protein termed HBZ. In light of the presence of potential ORFs in the antisense strand of HTLV-3 and HTLV-4, in this report, we were interested in evaluating if antisense transcription occurred in these viruses and whether it could encode a protein. Using proviral DNA constructs, RT-PCR analyses indeed indicated the existence of antisense transcription in transfected 293T cells. These transcripts were spliced and polyadenylated. The resulting protein termed APH (Antisense Protein of HTLV)-3 and APH-4 were both devoid of a bZIP domain but contained basic-rich regions. Confocal microscopy and Western blot experiments demonstrated a nuclear-restricted pattern for APH-4 while being both cytoplasmic and nuclear for APH-3. Promoter experiments also revealed that APH-3 and APH-4 were modulated by T cells activators. Both proteins were finally shown to be comparable in their capacity to inhibit Tax1- and Tax3-mediated HTLV-1 LTR activation. These results provide further evidence that antisense transcription in human retroviruses is not exclusive to HTLV-1 and that APH-3 and APH-4 likely impact on HTLV-3 and HTLV-4 replication, respectively.

3.3.2 Introduction

Human T-cell Leukemia Viruses (HTLV) are human deltaretroviruses, which are part of the Primate T-cell Lymphotropic Virus (PTLV) group and are constituted by two well studied members known as HTLV-1, the first retrovirus to be isolated in human (Miyoshi et al., 1981; Poiesz et al., 1980; Poiesz et al., 1981; Yoshida et al., 1982) and HTLV-2. HTLV-1 bears important significance in terms of human health as this virus is the etiological agent of Adult T-cell Leukemia/Lymphoma (ATLL) and HTLV-1-Associated Myelopathy/Tropical Spastic Paraparesis (HAM/TSP). Unlike HTLV-1, HTLV-2 has only been seriously linked to HAM-like pathologies (and not to leukemia). Recently, two new HTLV viruses, termed HTLV-3 and HTLV-4, have been isolated, the former having been closely compared to HTLV-3 (Calattini et al., 2006; Calattini et al., 2005; Wolfe et al., 2005). Currently, these viruses have been isolated in a relatively low number of individuals from Africa and no pathologies have yet been associated with these viruses.

The discovery of these two retroviruses has led to a series of recent studies comparing them to HTLV-1 and HTLV-2. These studies have indeed demonstrated that both HTLV-3 and HTLV-4 shared similar genomic organization to the former human retroviruses and weak but reproducible cross-reactivity with HTLV-1 and HTLV-2-specific antibodies were observed (Calattini et al., 2006; Calattini et al., 2005; Switzer et al., 2006; Wolfe et al., 2005). Further studies have focused on the Tax3 protein of HTLV-3 and have highlighted its similarity with HTLV-1 Tax in terms of both intracellular localization and transactivation activity (Calattini et al., 2006; Chevalier et al., 2006). A recent study has also provided evidence that the HTLV-3 genome, when reconstituted, expresses several different retroviral genes and produces infectious particles (Chevalier et al., 2008). Alike HTLV-1, these reports have also pointed toward the existence of a potential ORF, which could be produced from the antisense strand (Calattini et al., 2006; Chevalier et al., 2008; Switzer et al., 2006).

Former studies had indeed shed light on the existence of a new unexpected retroviral protein named HBZ (HTLV-I bZIP), which is encoded on the antisense strand of the HTLV-1 genome (Gaudray et al., 2002). Typically, two HBZ isoforms are produced with one of them being most abundant and dependent on a spliced transcript (Cavanagh et al., 2006; Murata et al., 2006; Satou et al., 2006). Both HBZ isoforms block Tax-induced and basal HTLV-1

transcription and interact with several Jun family members, rendering some of these members inactive through possible sequestration in transcriptionally inactive nuclear bodies (Basbous et al., 2003; Hivin et al., 2007; Matsumoto et al., 2005; Thebault et al., 2004). A number of reports have also demonstrated that HBZ was expressed in cells from ATLL patients and that it could therefore be implicated in ATLL development, partly through its hyperproliferative action on T cells (Arnold et al., 2008; Barbeau and Mesnard, 2007; Matsuoka and Jeang, 2007; Mesnard et al., 2006; Satou et al., 2006).

Although antisense transcription has been suggested in other retroviruses (Arnold et al., 2008; Briquet et al., 2001; Briquet and Vaquero, 2002; Michael et al., 1994; Peeters et al., 1996; Tagieva and Vaquero, 1997; Vanhee-Brossollet et al., 1995), more data are needed to clearly establish that indeed antisense transcription is present in several retroviruses beside HTLV-1. Given that former studies have alluded to the possible existence of antisense transcription in HTLV-3 and HTLV-4, we herein investigated on the presence of antisense transcripts from these viruses and on its potential encoding capacity. Our results indicate that both viruses produce a spliced and polyadenylated antisense transcript. The encoded proteins show distinct localization, the HTLV-3 antisense protein being both nuclear and cytoplasmic, while the HTLV-4 counterpart is almost exclusively contained in the nucleus. However, both antisense proteins were shown to inhibit Tax-mediated HTLV LTR activation. These results hence provide important data indicating the importance of antisense proteins in retroviral replication.

3.3.3 Material and methods

3.3.3.1 Cell lines and antibodies

The human and monkey kidney 293T and COS-7 cell lines were maintained in DMEM culture medium supplemented with 10% foetal bovine serum (Hyclone Laboratories). The T-cell line Jurkat E6.1 was maintained in RPMI-1640 medium supplemented with 10% foetal bovine serum (Hyclone Laboratories, Logan, UT), 2 mM glutamine, 100 U/ml penicillin G and 100 µg/ml streptomycin. The anti-Myc antibody (9E10) was purchased from Santa Cruz Biotechnology, the ECLTM sheep anti-mouse IgG antibody coupled to the horseradish peroxidase was obtained from GE Healthcare and the goat anti-mouse IgG antibody coupled to the Alexa fluor 488 (A11001) fluorochrome was obtained from Invitrogen.

3.3.3.2 Plasmids

Both HTLV-3 (2026ND) and HTLV-4 (pUC-HT4v2) proviral DNAs were cloned in pUC18. The pHTLV-3ΔEcoRV construct was derived from HTLV-3 (2026ND) proviral DNA by enzymatic digestion with EcoRV and XhoI, thereby deleting 4359pb of the 5' end of the proviral sequence. The pHTLV-3-as-luc vector was derived from pHTLV-3ΔEcoRV by cloning the luciferase reporter gene from pGL3-basic in frame in the second exon of APH-3 at position 7181 (sense orientation). Briefly, PacI and BstZ171 restriction sites were added to the pHTLV-3ΔEcoRV construct by PCR using PhusionTM Site-Directed Mutagenesis Kit according to manufacturer's instructions using the reverse primer 26-1 (5'-P-ATAG**TATAC** TGCAATCCCAGGAACTG-3') containing BstZ171 restriction site (bold) and forward primer 27-1 (5'-P-ATAT**TAATTA**ATGTCTCCGGGGCTAGG-3') containing PacI restriction site (bold). Luciferase was amplified using reverse primer 35-1 (5'-ATAT**TAATTA**AAGGAAGACGCCAAAAACATAAAGAA-3') containing PacI restriction site (bold) and forward primer 32-1 (5'-

ATAGTATACTACCACATTTGTAGAGGTTTTAG-3') containing BstZ171 restriction site (bold). Luciferase amplicons were digested by PacI and BstZ171 and ligated in the pHTLV-3ΔEcoRV digested vector. The pHTLV-4ΔSacI construct was derived from HTLV-4 (pUC-HT4v2) proviral DNA by enzymatic digestion with SacI and HindIII, thereby deleting 3213pb of the 5' end of the proviral sequence. The pHTLV-4-as-luc vector was obtained by cloning de luciferase reporter gene in frame in the second exon of APH-4 at position 7037 (sense transcript) in pHTLV-4Δ6560 (construct containing the last 2183bp of the HTLV-4 genome in pUC18 vector). Briefly, PacI and BstZ171 restriction sites were added to the pHTLV-4Δ6560 construct by PCR using PhusionTM Site-Directed Mutagenesis Kit according to manufacturer's instructions using the reverse primer 26-2 (5'-**ATAGTATACAGAGGAGATGCCTGGTA-3'**) containing BstZ171 restriction site (bold) and forward primer 31-1 (5'-**ATATTAATTAATGGTGTGAGACCTTCTTTG-3'**) containing PacI restriction site (bold). Luciferase was amplified using reverse primer 35-1 (5'- **ATATTAATTAAGGAAGACGCCAAAAACATAAAGAA-3'**) containing PacI restriction site (bold) and forward primer 32-2 (5'-**ATAGTATACTACCACATTTGTAGAGGTTTTAG-3'**) containing BstZ171 restriction site (bold). Luciferase amplicons were digested by PacI and BstZ171 and ligated in the pHTLV-4Δ6560 digested vector. The pHTLV-Luc vector contains the complete HTLV-1 LTR 3' region cloned into the XhoI/HindIII site of the pGL2-Basic vector (Promega) (Geleziunas et al., 1998). The Tax1 expressing vector has been previously described (Matsumoto et al., 1997). The pCMV-Tax3 and pCMV-Tax4 vectors contain the Tax3- and Tax4-coding regions under the control of the CMV promoter region in the pcDNA3.1Zeo (+). The pMycAPH-3 and pMycAPH-4 were generated by cloning the RT-PCR amplicon of the coding regions of APH-3 or APH-4 using a primer that contain a Myc tag in the pcDNA3.1Zeo (+) vector. Briefly, APH-3 was amplified using reverse primer 53-1 (5'-**ATGGAACAAAACTCATCTCAGAAGAGGATCTGATGGCTCGATCCCGAAGCGG-3'**) containing a Myc tag (bold) and forward primer 36-1 (5'-**ATATCTAGATTATAACAGATCTGCTACCTCCTGTAG-3'**) containing a XbaI restriction site (bold). APH-4 was amplified using reverse primer 59-1 (5'-**ATGGAACAAAACTCATCTCAGAAGAGGATCTGATGGACACTCGAGAATTTTTTAGGGG-3'**) containing a Myc tag (bold) and forward primer 37-1 (5'-

ATATCTAGATTATAATAACTCCGCCAATACACCCAAC-3') containing a XbaI restriction site (bold). Amplified products were digested by XbaI and ligated into pcDNA3.1Zeo (+) digested by XbaI and EcoRV. The pRcActin-LacZ vector contains the β -galactosidase gene under the control of the β -actin promoter.

3.3.3.3 Transfection

293T cells were transfected with 10-15 μ g of DNA using the calcium phosphate protocol as previously described (Fortin et al., 1997). For the luciferase experiment, 293T cells were transiently transfected using the LipofectamineTM Reagent (Invitrogen) according to manufacturer's instructions with 0,4 μ g of pHTLV-luc vector and 0,2 μ g of pCMVTax3 or pCMVTax4 together with 0,2 μ g of the expression vectors for pMycAPH-3, pMycAPH-4 or HBZ-SP1-Myc (compared to the empty vector pcDNA3.1) along with 0,1 μ g of the pRcActin-lacZ vector for normalization. 293T transfected cells were lysed 48h post-transfection in lysis buffer (25mM Tris phosphate, pH 7.8, 2mM DTT, 1% Triton X-100, 10% glycerol) and luciferase activity readout was performed with the MLX microplate luminometer (Dynex Technologies) with a single injection of a luciferase buffer (20 mM tricine, 1.07 mM (MgCO₃)₄·Mg(OH)₂·5H₂O, 2.67 mM MgSO₄, 0.1 mM EDTA, 220 μ M Coenzyme A, 4.7 μ M D-Luciferin potassium salt, 530 μ M ATP, 33.3 mM DTT). Each sample was co-transfected with a β -gal-expressing vector for transfection efficiency normalisation. The β -galactosidase activity was measured using the Galacto-LightTM kit (Applied Biosystems, Bedford, MS) according to manufacturer's suggestions. Luciferase activity are presented in normalized Relative Light Units (RLU/ β -gal) and represent the calculated mean \pm SD of three transfected samples normalised by the measured β -galactosidase activity. Jurkat cells were transfected by electroporation with 15 μ g of DNA using the Gene Pulser Xcell system (BioRad, Hercules, CA) (960 μ F, 250V). Jurkat cells were resuspended at 2,5 \times 10⁷/ml of complete medium, transfected in bulk and were separated at 16h post-transfection into various treatment groups at a density of 1 \times 10⁵cells/well (100 μ l) in 96-well flat-bottom plates. Cells were either left untreated or were treated with phorbol 12-myristate 13-acetate (PMA) at 20 ng/ml (Sigma-Aldrich, Oakville, Canada),

phytohemagglutinin (PHA-P) at 3 $\mu\text{g/ml}$ (Sigma-Aldrich, Oakville, Canada), ionomycin (Sigma-Aldrich, Oakville, Canada) at 1 μM , bpV [pic] (Alexis corp.) at 15 μM , forskolin (Sigma-Aldrich, Oakville, Canada) at 10 μM and $\text{TNF}\alpha$ at 20 ng/ml. Luciferase activity was monitored at 8h post-stimulation.

3.3.3.4 RT-PCR and 3'RACE analyses

Total RNA was extracted by the TRIzol reagent (Invitrogen) from transfected 293T cells 48 h post-transfection. PolyA⁺ RNA was purified from total RNA using Oligotex mRNA Mini Kit (Qiagen) according to manufacturer's instructions. RT-PCR analyses were conducted using an oligo(dT) primer (Invitrogen). Briefly, total RNA (5 μg) was mixed with 1 μl of 10 μM of the oligo(dT) primer. The RNA:RT primer mix was heated at 70°C for 5 min, chilled on ice for 5 min and incubated 2 hours at 42°C in the presence of 1X AMV reaction buffer, 1 mM dNTPs, 10U of SUPERase-In RNase inhibitor (Ambion) and 15 U AMV reverse transcriptase (USB). Synthesized cDNAs were then PCR amplified in the presence of 1.25 U Taq DNA polymerase, 1X ThermoPol buffer, 20 μM dNTPs, 1,5 μM of each primer using a Tgradient thermocycler (Biometra). Primers used for PCR reactions were the reverse primer 23-1 (5'-CCACCAAATAGAAGAGGGATGGC-3') and forward primers 19-1 (5'-GACGCCCTGGCCCCAACAG-3') for the amplification of the antisense transcript of HTLV-3 and the reverse primer 21-1 (5'-CGGCGGCGTCTCAACTGATTG-3') and forward primer 23-4 (5'-ACGAGTCCCCCATATGTCCAAA-3') for the amplification of the antisense transcript of HTLV-4. PCR conditions were as follow: a first step of denaturation at 94°C for 5 min followed by 35 cycles of denaturation (94°C for 1 min), annealing (60°C for 1 min) and extension (72°C for 1 min) and a final extension at 72°C for 5 min. RT-PCR amplification were controlled for DNA contamination (RNA samples with no RT step). The polyA signals were identified from isolated polyA⁺ RNA with the FirstChoice RLM-RACE kit from Ambion according to manufacturer's instructions. The cDNA synthesis was performed in the presence of the supplied 3'RACE adapter. PCR amplification was achieved through 3'RACE outer and inner primers and a primer derived from the sequence downstream of the ORFs: 23-2 (5'-GGAGAGGAACCACACTGGATCAT-3') for HTLV-3

antisense transcript and 23-5 (5'-GAGTCAGGACATGCTCTAGGTCT-3') for HTLV-4 antisense transcript. Amplified products were then cloned in pBlueScript KS and sequenced.

3.3.3.5 Confocal microscopy

COS-7 or 293T cells were seeded onto a 6 well plate containing a 1.5mm thick coverslip for 24 h before transfection and were then transfected using the Lipofectamine™ Reagent (Invitrogen) according to manufacturer's instructions. Cells were analyzed by confocal microscopy at 36h post-transfection. Briefly, cells were washed with PBS, fixed with 4% paraformaldehyde for 15 min and permeabilized with 0.1% Triton X-100 for 5 min at room temperature. Cells were then washed 3 times with PBS and incubated with a mouse anti-Myc 9E10 antibody (dilution 1:400) for 1h at room temperature. Samples were next washed 3 times with PBS and incubated with a secondary goat anti-mouse IgG antibody coupled to a green-fluorescent Alexa fluor 488 (Invitrogen) for 45min at room temperature and washed again with PBS. Cells were then incubated in a 2,5 µg/ml propidium iodide (PI) solution for nuclear DNA staining. The coverslip were then mounted in a drop of ProLong Antifade (Invitrogen). Images were acquired with a MRC1024 confocal laser-scanning microscope (BioRad, Hercules, CA).

3.3.3.6 Western blot analysis

Cytoplasmic and nuclear extracts from transfected 293T cells were prepared using NE-PER® Nuclear and Cytoplasmic Extraction Reagents (Pierce Biotechnology) according to manufacturer's instructions. Equal quantities of extracts were run on a SDS-10% PAGE and transferred to PVDF membranes (Millipore). The blot was next blocked in PBS/5% milk and incubated with a mouse anti-Myc 9E10 antibody (dilution 1:1000). After several washes, signals were revealed by the addition of peroxylase-conjugated goat anti-mouse IgG (dilution 1:10000) antibody and subsequent incubation with the BM Chemiluminescence Blotting Substrate (POD) (Roche). Membranes were exposed on Amersham hyperfilms ECL (Amersham Pharmacia Biotech).

3.3.4 Results

3.3.4.1 Detection of spliced antisense transcripts in HTLV-3 and HTLV-4-transfected cells.

Previous studies have highlighted the presence of a potential ORF in the antisense strand of HTLV-3 (Calattini et al., 2006; Chevalier et al., 2008; Switzer et al., 2006). Given that the existence of the antisense protein has been exclusively demonstrated in this retrovirus family for HTLV-1, we therefore asked whether both newly discovered HTLV-3 and HTLV-4 human retroviruses could express similar antisense protein.

Our first *in silico* analysis of the ORF present in both previously described HTLV-3 and HTLV-4 proviral DNA (Switzer et al., 2006; Wolfe et al., 2005) indeed revealed the existence of an ORF in the antisense strand between both *tax* and *env* genes (Figure 3.1A and B). The encoded protein was predicted to contain a maximum of 221 and 158 amino acids (for HTLV-3 and HTLV-4 respectively) from a predicted ATG initiation codon. These ORFs were termed APH (Antisense Protein of HTLV)-3 and APH-4.

To verify the existence of an antisense transcript from these retroviruses, we first conducted RT-PCR analyses. Each proviral DNA constructs were first deleted of their 5'LTR region to minimize interference from sense transcription (Figure 3.1). The resulting pHTLV-3ΔEcoRV and pHTLV-4ΔSacI proviral DNA constructs were next transfected in 293T cells and analysed by RT-PCR using forward and reverse primers derived from the 3'LTR and APH-3/APH-4 regions, respectively (see Figure 3.1A and B). The choice of primers and their positioning were based on the previously described spliced pattern of the antisense HBZ transcript of the related HTLV-1 virus (Cavanagh et al., 2006; Murata et al., 2006; Satou et al., 2006). RT-PCR analyses indeed revealed the presence of a single signal specific for an antisense transcript (as demonstrated by the absence of a signal in the absence of RT). In addition, its size also lends credence to the occurrence of splicing while no signals with a size expected for unspliced transcripts were amplified (Figure 3.2A and B). Splice sites of the HTLV-3 antisense transcript were positioned at nt 8556 and 7218 for the splice donor and splice acceptor, respectively, while their corresponding SD and SA sites were at nt 8422 and

7089 for the HTLV-4 antisense transcript. Position of the splice sites and the length of the intronic region were very well conserved between HTLV-3 and HTLV-4. Similarly to HTLV-1 HBZ, in both APH-3 and APH-4 transcripts, the resulting spliced transcript led to the addition of an *N*-terminal amino acid sequence with a methionine initiation codon to the remaining corresponding APH ORF amino acid (Figure 3.2C). Interestingly, for APH-3, a total of 9 amino acids were added from exon 1 while APH-4 exon 1 was responsible for the addition of 4 amino acids. An important observation is that unlike HBZ, if the unspliced mRNA could encode for a different APH isoform, in both cases, the resulting encoded APH isoform would initiate at a position downstream of the splice acceptor site thereby leading to a shorter isoform than the protein potentially encoded from a spliced transcript.

These results hence demonstrated that both viruses were capable of producing a spliced antisense transcript which could produce a protein. Signals corresponding to an unspliced transcript were not detected in these analyses.

3.3.4.2 *APH-3* and *APH-4* transcripts are both polyadenylated.

We were next interested in characterizing the 3' end of both retroviral antisense transcripts (Figure 3.3). Again, proviral DNA-deleted versions of HTLV-3 and HTLV-4 were transfected in 293T cells and RNA were used to identify the 3' end of the transcripts by 3' RACE analysis. As depicted in Figure 3.3, both transcripts were cleaved and polyadenylated between a consensus polyA signal and a GT-rich sequence, which is often found in proximity to the polyA tail addition site. Interestingly, although the composition of the targeted dinucleotide for cleavage was different for both retroviruses (AG and CC for HTLV-3 and HTLV-4, respectively), both cleavages occurred at an equal distance from the polyA signal (19 nt).

3.3.4.3 Amino acid sequence of APH-3 and APH-4 and their cellular localization.

The resulting amino acid sequence of APH-3 and APH-4 were compared to the previously published HBZ sequence (Figure 3.4). This comparison revealed that the

previously described bZIP domains of HBZ was not conserved among both APH-3 and APH-4. Basic rich regions were however noted at similar positions for all three positions and could reflect potential NLS sequence for both APH-3 and APH-4. In addition, similar LXXLL and LXXLL-like motifs known to be responsible for an interaction of HBZ with CBP/p300 (Clerc et al., 2008) were identified in both APH-3 and APH-4 proteins.

Based on these analyses, we next determined the cellular localization of APH-3 and APH-4 using expression vectors for both proteins tagged with a Myc epitope at their amino ends. These vectors were both transfected in 293T and COS-7 cells and analyzed by confocal microscopy. Our first analysis of APH-3 revealed that it localized mainly in the nucleus but showed significant staining in the cytoplasm (Figure 3.5A). This staining pattern could be observed from 24 to 72 hours post-transfection (data not shown). In addition, punctate structures were observed in stained COS-7 nuclei. Upon merging, nuclear signals were however shown to poorly co-localize with the nucleolus.

Similar analyses were conducted on 293T and COS-7 cells transfected with the Myc-tagged APH-4 expression vector. Interestingly as opposed to APH-3 and more reminiscent of HBZ, APH-4 was strictly present in the nucleus but again did not co-localize with the nucleolus (Figure 3.5B). All these confocal microscopy experiments were retested on cells transfected with COOH-tagged APH expression vectors and again led to similar results (data not shown) indicating that the added tag and its position likely did not contribute to the cellular distribution of APH-3 and APH-4.

To confirm these results, Western blot analyses were performed on nuclear vs. cytoplasmic extracts prepared from 293T cells transfected with expression vectors for Myc-tagged HBZ, APH-3 or APH-4 (Figure 3.5C). Controls consisted of extracts from cells transfected with the empty expression vector. Extracts indeed revealed the expected size for APH-3 (28 kDa) while APH-4 indicated a lower molecular weight than the predicted size being 22 kDa. As previously reported (Arnold et al., 2006; Cavanagh et al., 2006), HBZ showed a slower migration than the expected size of 27 kDa. In addition, as predicted, both APH-4 and HBZ were mainly present in nuclear extracts while showing a weak presence in cytoplasmic extracts. On the other hand, the APH-3 signal was present in both extracts. Interestingly, a higher molecular weight signal was apparent exclusively in the nucleus of

APH-3-expressing cells although the signal might be non-specific as judged by the presence of a weaker bands in control transfected cells.

These results hence demonstrated that both APH-3 and APH-4 are differently distributed in transfected cell, which might impact on their function in their respective viruses.

3.3.4.4 Promoter studies on both HTLV-3 and HTLV-4 antisense transcripts.

To further analyse the *APH-3* and *APH-4* genes, expression levels were analysed in Jurkat T cells using constructs in which a luciferase reporter gene along with the poly A signal was inserted in frame with the amino acid sequence present in exon 2 (Figure 3.6A). Jurkat cells were subsequently transfected and then either left untreated or stimulated with a series of known T-cell activators. In non-stimulated conditions, luciferase activity was found to be modest but significant showing no strong difference between HTLV-3 and HTLV-4 constructs (data not shown). Stimulation also led to an increase in luciferase activity for both vectors and were more responsive to the addition of the PTP inhibitor, bpV[pic], a strong T-cell activating agent.

3.3.4.5 Tax-driven HTLV LTR activation is inhibited by APH-3 and APH-4 expression.

Former studies have demonstrated that the HTLV-I-encoded antisense protein HBZ inhibited the Tax-mediated activation of its promoter. In addition, this inhibition was mediated by direct interaction with CREB-2 through its leucine zipper domain. As APH-3 and APH-4 seems to lack such a domain, we asked whether it could nonetheless act upon Tax-mediated HTLV-1 LTR activation. Expression vectors for APH-3, APH-4 and HBZ (in comparison to pcDNA3.1Zeo (+)) were transfected in 293T cells along with expression vectors for Tax 1 or Tax3 and pHTLV-luc. Measurement of normalized luciferase activity demonstrated that, as previously reported, HBZ blocked Tax activation of the HTLV-1 LTR.

Importantly, both APH-3 and APH-4 showed a similar capacity to block this Tax-mediated HTLV-1 LTR upregulation, albeit at different levels.

These results therefore indicated that, although APH-3 and APH-4 lack a typical bZIP domain, both proteins can alter the extent of activation of the HTLV-1 5'LTR mediated by either Tax1 or Tax3 and could therefore play a similar role in HTLV-3 and HTLV-4 replication.

3.3.5 Discussion

HTLV-3 and HTLV-4 retroviruses have been recently isolated from Cameroon individuals, which had no sign of illness (Calattini et al., 2005; Wolfe et al., 2005). Studies have further demonstrated that a number of retroviral genes are shared with the related HTLV-1 and HTLV-2 viruses (Calattini et al., 2006; Chevalier et al., 2006; Switzer et al., 2006). Given that recent evidence points to the importance of the newly discovered antisense-encoded HBZ gene in both HTLV-1 replication and ATL development, the goal of this study was to provide evidence for the existence of antisense transcripts in these viruses and to further assess the possible functional relevance of their encoded proteins.

Our first *in silico* analyses confirmed previous analyses in that a potential ORF was indeed being present on the antisense strand of the HTLV-3 and HTLV-4 genomes. Our analyses have indeed demonstrated that these proteins were bZIP-less but contained a certain number of basic-rich regions similar to basic domains identified for HBZ. Although these basic amino acid-rich regions could serve in targeting APH-3 and APH-4 to the nucleus, these domains could also participate in modulating the transcriptional modulation of APH-3 and APH-4.

Our RT-PCR analyses and 3' RACE experiments have conclusively indicated that the antisense transcripts of HTLV-3 and HTLV-4 are spliced and polyadenylated. The splicing pattern was similar to that of HBZ with a similar intronic size. Although more analyses are needed, current results argue for low abundance of unspliced transcripts and therefore suggest that no other protein isoform could be produced from the antisense transcript. In addition, unlike HBZ, translation of APH-3 and APH-4 would initiate with a downstream methionine resulting in the deletion of 14 and 24 amino acids from the NH₂ end, respectively. Such deletions might cause important change in the protein and might suggest why unspliced transcripts are not abundant and might therefore not lead to synthesis of another isoform.

Sequence analyses of the SD and SA sequence have revealed typical consensus sequence. Consensus poly A signal and GT-rich sequence were also identified in proximity to the poly A addition site for both HTLV-3 and HTLV-4. This is again reminiscent of the situation of the HBZ transcript. In fact positioning of the poly A addition site for *APH-3* and

APH-4 RNA is equivalent to the HBZ gene and likely reveals specific constraints associated with antisense transcription.

The analysis of the amino acid composition of both APH-3 and APH-4 and their comparison to HBZ demonstrated important differences. First and importantly, software analysis of APH-3 and APH-4 did not highlight a LZ domain, a marked difference with HBZ. In addition, APH-3 showed important differences in amino acid sequence. However, it should be stressed that like HBZ, basic regions were identified and LXXLL and LXXLL-like motifs (known to be responsible for binding p300/CBP) were also noted. These similarities might indeed contribute to the similarity in cellular localisation and functional aspects between APH-3, APH-4 and HBZ. Confocal and Western blot experiments have indeed indicated that APH-3 and APH-4 are both nuclear (at least partly for APH-3). The NLS sequence which mediates nuclear targeting remains however to be determined but might reasonably involve similar regions to those responsible for nuclear localisation of HBZ. However, unlike HBZ (the SP1 isoform), neither APH-3 nor APH-4 co-localize with the nucleolus. In fact, their nuclear distributions are more representative of the formerly published localisation of the HBZ isoform, thought to be produced from unspliced transcripts (Gaudray et al., 2002). Further experiments will be required to determine how these differences in nuclear distribution could be explained.

Another interesting observation is the cytoplasmic localization of APH-3 observed in both 293T and COS-7 cells. This observation might be indicative of the existence of a potential nuclear export signal. It is also important to note that a higher molecular weight signal was noted in our Western blot analysis in APH-3-expressing cells and that this signal was mainly localised to the nucleus. This opens up the possibility that nuclear localisation for APH-3 (and even for HBZ and APH-4) could be importantly modulated by potential post-translational modifications, such as phosphorylation or sumoylation. Interestingly, the viral Tax protein has been previously shown to be altered in its nuclear distribution by sumoylation events (Lamsoul et al., 2005). Further experiments will be required to assess if this is indeed the case.

Studies of the APH-3 and APH-4 genes using the luciferase reporter gene have further showed that their expression was stimulated by known T-cell activators. Similar results have been obtained with constructs similar for the HBZ gene. It might be expected

that as for HBZ expression, Tax expression would induce its expression *via* the equivalent TRE sequences present in the U3 region (Landry et al., 2008; Yoshida et al., 2008).

Functional analyses have been investigated for both APH-3 and APH-4. HTLV-1 LTR has been previously shown to also be activated by Tax3 (Calattini et al., 2006). Our analyses have indeed revealed that both antisense proteins could indeed inhibit Tax1 and Tax3 activation. This raises the issue, as we have previously demonstrated, that the leucine zipper domain is not the only amino acid segment responsible for Tax inhibition by HBZ (Clerc et al., 2008). Amino acid sequence of both APH-3 and APH-4 highlights possible LXXLL and LXXLL-like motifs which are known to mediate interaction with p300/CBP and are related to the interaction of HBZ and p300/CBP. Hence, the mechanism of Tax inhibition is likely different from HBZ and should further be considered in the context of the differences in cellular distribution of all three antisense proteins. These differences could also impact on the cellular genes differently modulated by HBZ, APH-3 and APH-4 and might affect their capacity to modify cellular proliferation or transformation.

The current data demonstrate that antisense transcription is a common mode of expression in HTLV viruses (and likely STLV family members). Recent results have also confirmed that a similar antisense transcript and encoded protein was being produced from the HTLV-2 virus (Halin et al., 2008), a virus which has not been related to any haematological malignancies. The similarity between these transcription patterns and the function of their encoded proteins highlights the importance of antisense-encoded HTLV gene for replication. Furthermore, a link between HBZ and ATL development has been suggested. It is not currently known if HTLV-3 and HTLV-4 are also associated with human diseases and future studies will help in assessing this possibility as well as the possible association of viral proteins (including APH-3 and APH-4). In addition, it will be interesting to determine if other human and non-human retroviruses might also be capable of producing antisense transcripts with an encoding capacity.

3.3.6 Acknowledgments

This work was supported by a grant to B.B. from the Cancer Research Society (CRS) Inc. M.H. is supported by an institutional Hydro-Quebec scholarship and S.L. holds a CIHR PhD scholarship. B.B. holds a Canada Research Chair of Canada (Tier 2).

3.3.7 Figure legends

Figure 3.1 : Position of the antisense ORF in the HTLV-3 and HTLV-4 genomes.

(A) HTLV-3 2026ND (A) and HTLV-4 HT4v2 (B) molecular clones are depicted with all known genes. The positioning of the putative APH-3 and APH-4 ORF from the antisense strand is indicated below proviral DNA. A 5' deleted version of these vectors termed pHTLV-3 Δ EcoRV and pHTLV-4 Δ SacI are also presented. The size of the proviral DNA (full length or 5' end deleted version) are indicated for each retrovirus. The primers used for the RT-PCR experiments are also represented.

Figure 3.2 : HTLV-3 and HTLV-4 antisense transcripts are detected and are spliced.

293T cells were transfected either with pHTLV-3 Δ EcoRV (A) or pHTLV-4 Δ SacI (B). RNA was extracted and analyzed by RT-PCR for the presence of antisense transcripts using the primer combinations 21-3/19-1 (A) or 21-1/23-2 (B). Deduced spliced transcript for each proviral DNA are depicted and position of the SD and SA sites are indicated below the spliced transcript (nucleotidic positioning from sense strand). Control (-) = no cDNA added in the PCR reaction; M= 100 bp marker. (C) Amino acid sequence of the spliced *APH-3* and *APH-4* RNA next to the spliced junction. The amino acid is shown above each sequence. As a reference, the nucleotidic and amino acid sequences predicted for unspliced transcripts are also provided for comparison.

Figure 3.3 : Identification of the poly A addition site for both *APH-3* and *APH-4* transcripts.

RNA samples obtained from 293T cells transfected with either pHTLV-3 Δ EcoRV (A) or pHTLV-4 Δ SacI (B) were analyzed by 3' RLM-RACE. Lane (-): PCR amplification in the absence of any cDNA. M: 100 bp marker. The sequences of the *APH* mRNA and of the 3'

poly A tail are shown. In the sequence below, the position of the cleavage site and the presence of both poly A signal and GT-rich sequence are highlighted.

Figure 3.4 : Amino acid sequence of the APH-3 and APH-4 proteins.

The amino acid sequence of APH-3 and APH-4 were deduced from the 2026ND and HT4v2 strains, respectively. Both sequences were compared to the amino acid sequence of the HBZ-SP1 isoform. The previously described BR2, BR1 and DNA-binding domain (DBD) are indicated above the HBZ sequence, while the LZ domain is in a box.

Figure 3.5 : APH-3 and APH-4 demonstrate different cellular localizations.

(A) pMycAPH-3 (B) and pMycAPH-4 expression vectors were transfected into 293T and COS-7 cells. Thirty six hours post-transfection, cells were fixed and stained as described in Material and Methods section. Cells were finally mounted in ProLong Antifade in the presence of propidium iodide. Images are representative of the entire population of transfected cells. Images on the right side of the panels are presented after merging of both anti-Myc and PI signals. (C) Cell lysates (50 μ g) from 293T cells transfected as in (A-B) or with the pHBZ-Myc vector were first separated in cytoplasmic (Cyt) and nuclear (Nuc) fractions and were subjected to electrophoresis on a 12 % Bis-Tris gel and analyzed by Western blot with anti-Myc, antibodies. pcDNA3.1/Myc parental vector was used as a negative control (CTL). For APH-3, the asterisk points to the major signal with the expected size.

Figure 3.6 : Antisense expression is upregulated in Jurkat T cells after activation.

(A) Luciferase reporter gene was inserted in frame with the exon 2 encoding segment of both APH-3 and APH-4 along with a poly A signal cassette at its 3'end generating pHTLV-3-as-luc and pHTLV-4-as-luc. (B) Both luciferase-encoding constructs were transfected in Jurkat T cells and were then activated for 8 hours with PHA, PMA, ionomycin, bpV[pic], forskolin, TNF- α or a combination. Luciferase activity was measured from cell lysates of three

independently stimulated samples and are presented as the average fold induction (\pm S.D.) where a value of 1 is given to unstimulated cells. These results are representative of four independent experiments.

Figure 3.7 : APH-3 and APH-4 repressed Tax1 and Tax3 dependent transcription from the HTLV-1 LTR.

293T cells were transiently transfected with pHTLV-luc (left) and pCMVTax3 (upper panel) or pCMVTax4 (bottom panel) together with expression vectors for pMycAPH-3, pMycAPH-4 or HBZ-SP1-Myc (compared to the empty vector pcDNA3.1) along with the pRcActin-lacZ vector for normalization. The presented luciferase activities are an average of 3 independent transfection experiments and are depicted as the average normalized luciferase activity \pm S.D. These results are representative of three independent experiments.

3.3.8 Figures

A

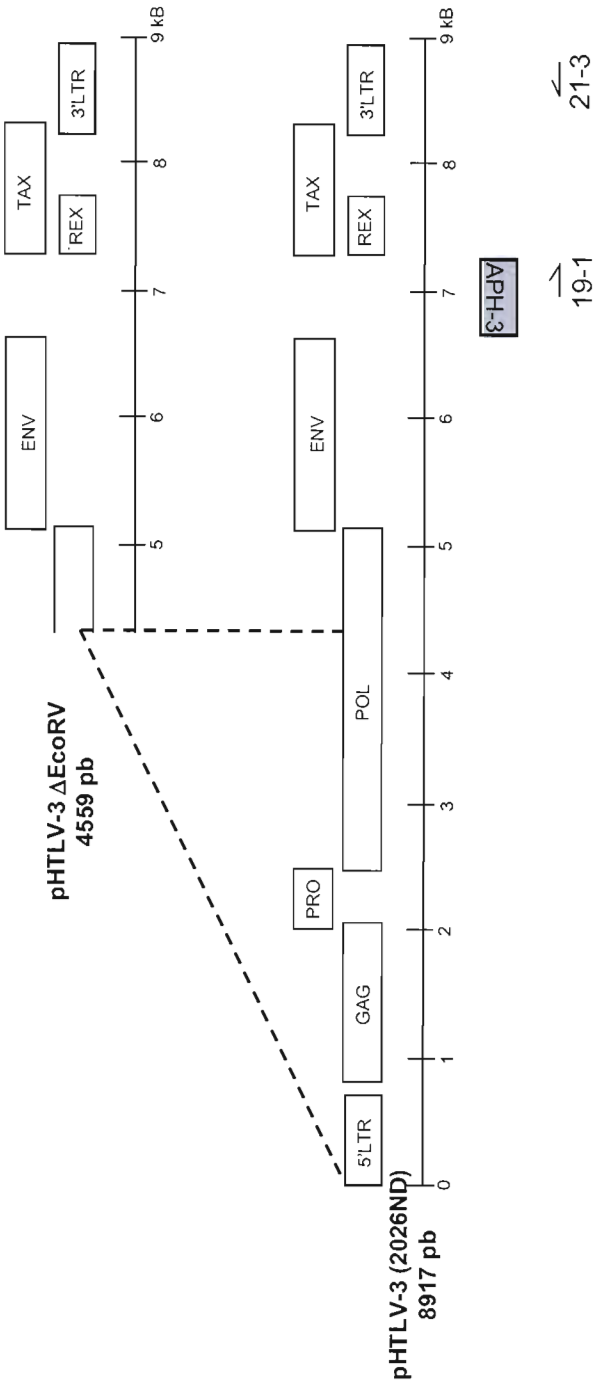


Figure 3.1

B

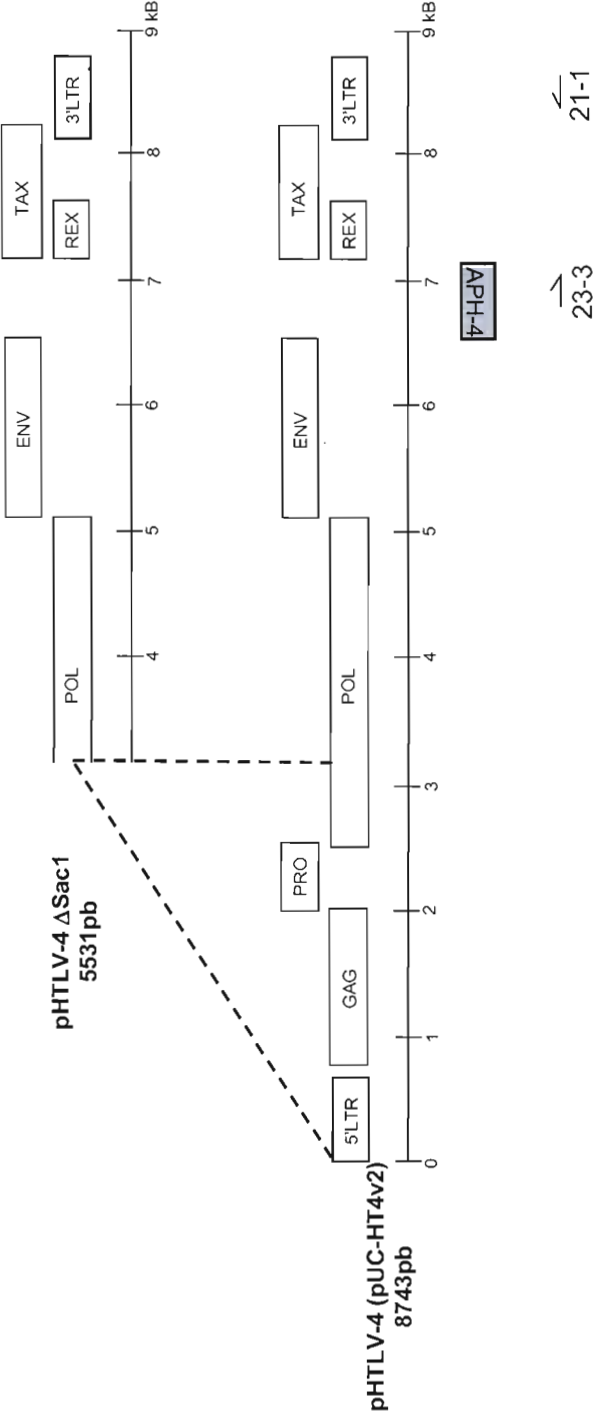


Figure 3.1

A

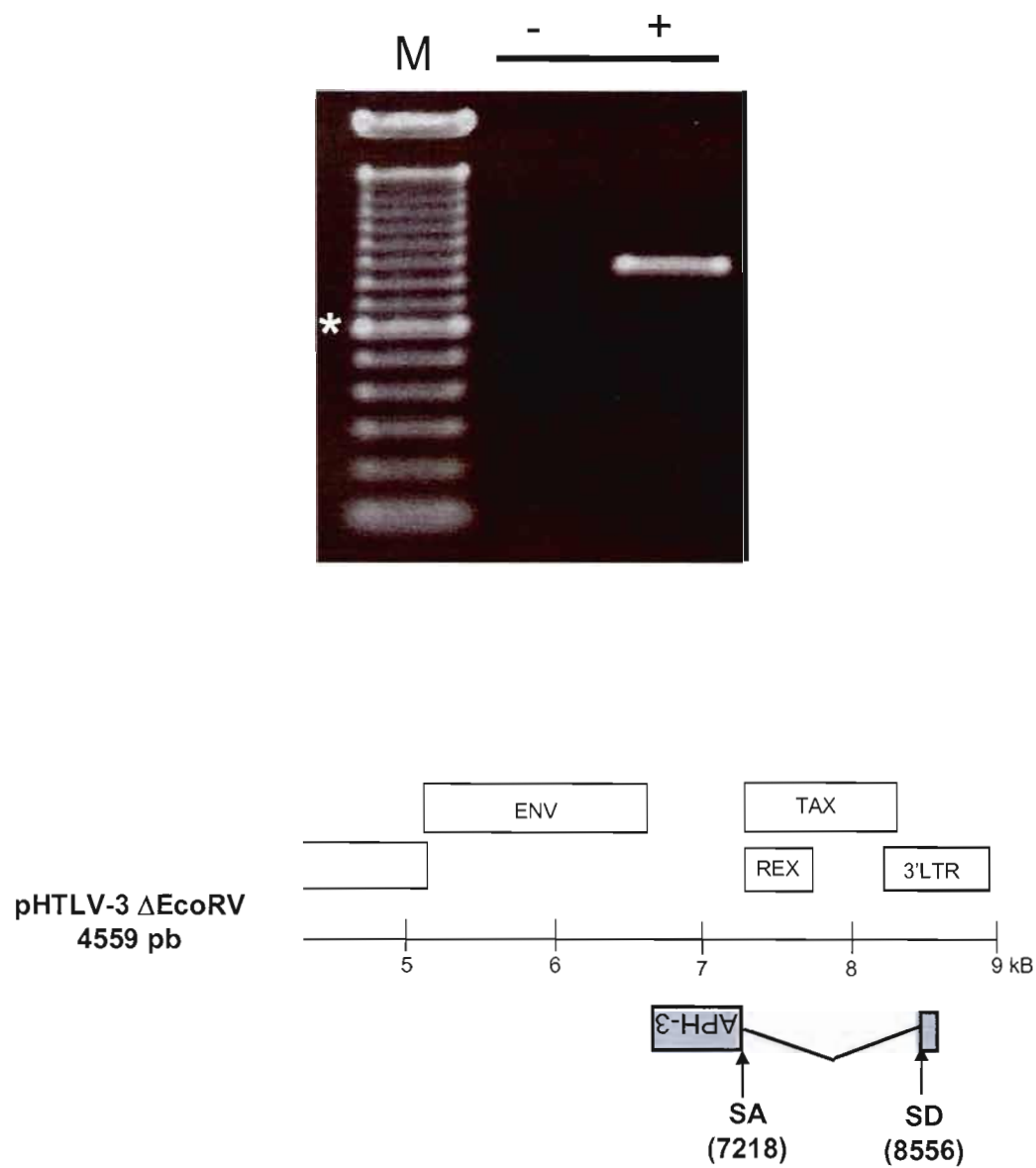


Figure 3.2

B

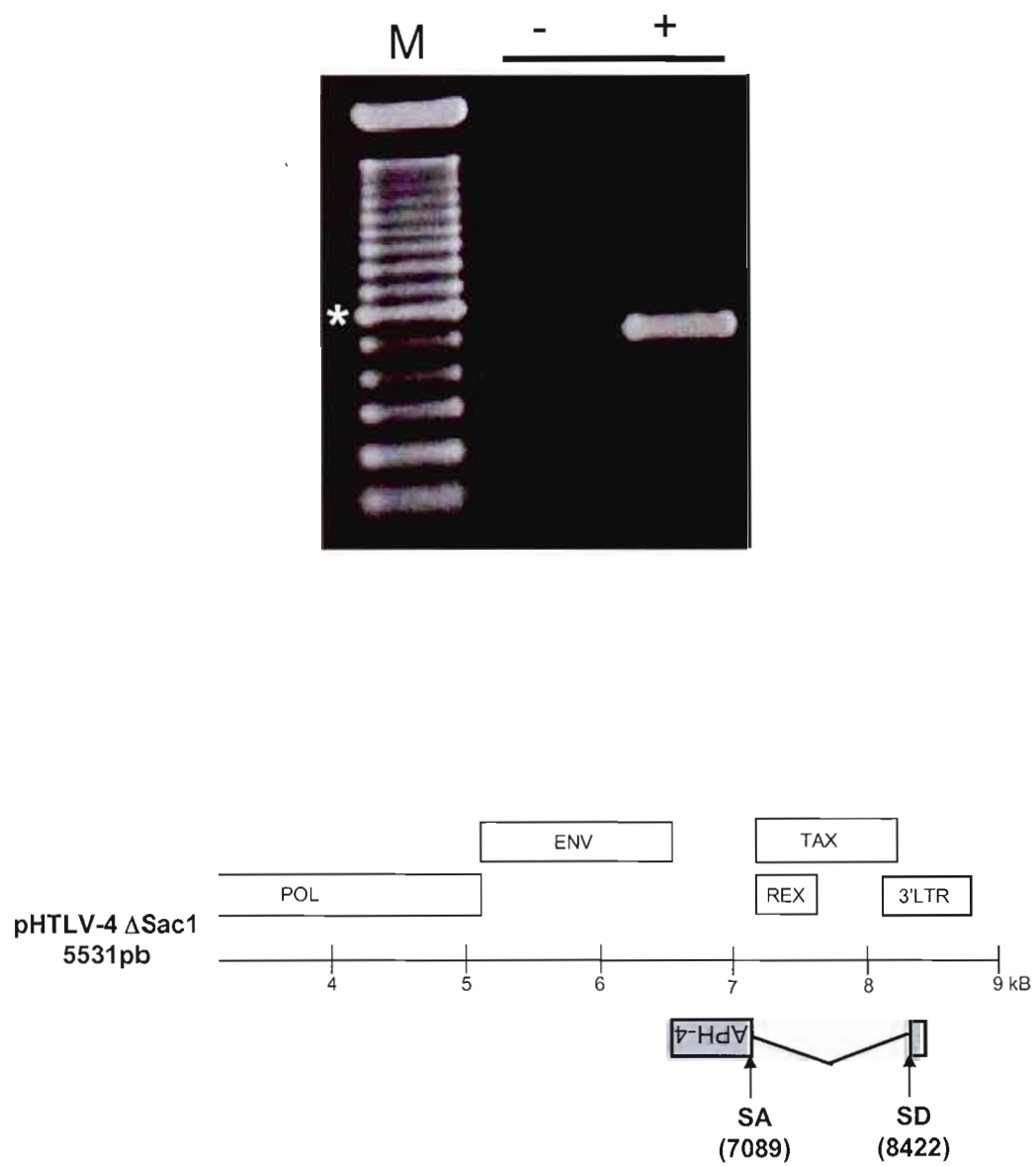


Figure 3.2

C

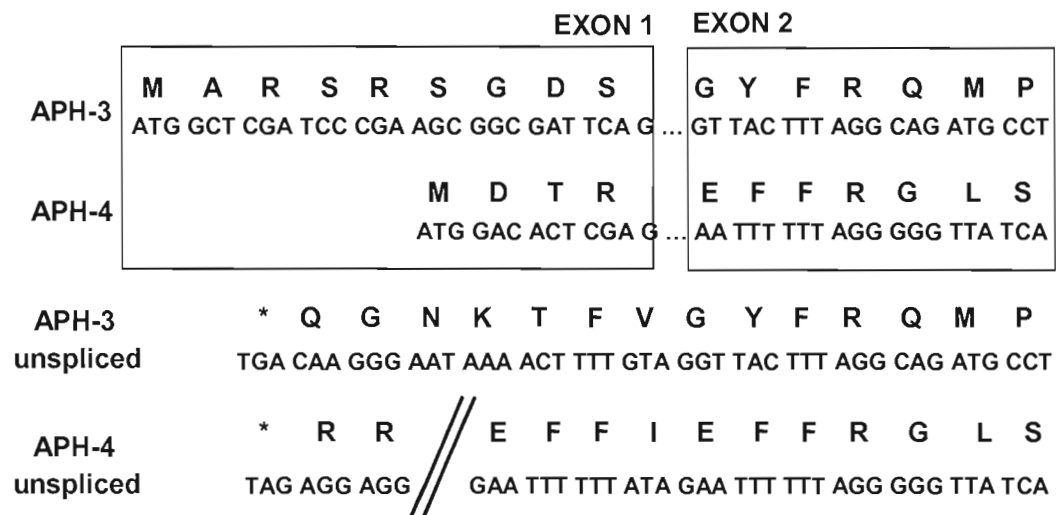


Figure 3.2

A

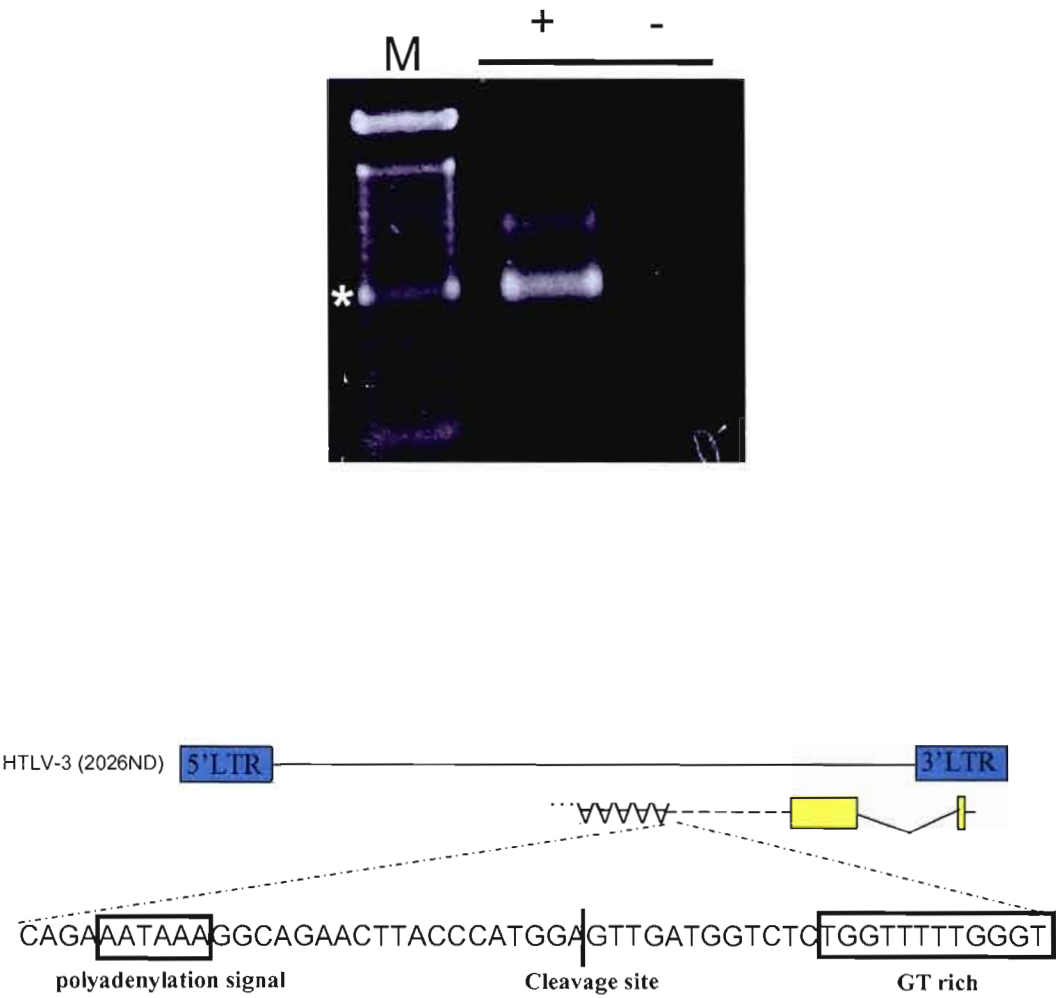


Figure 3.3

B

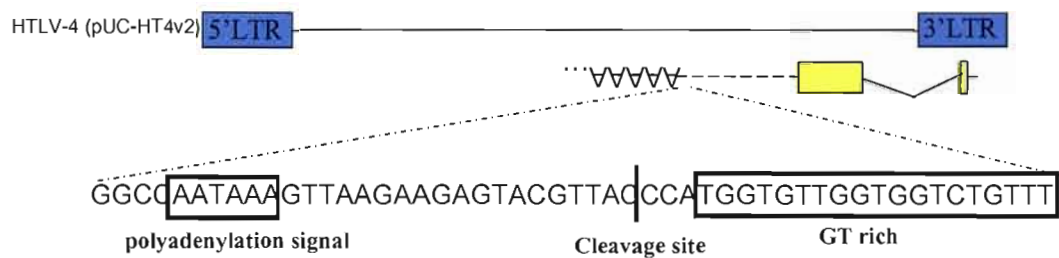
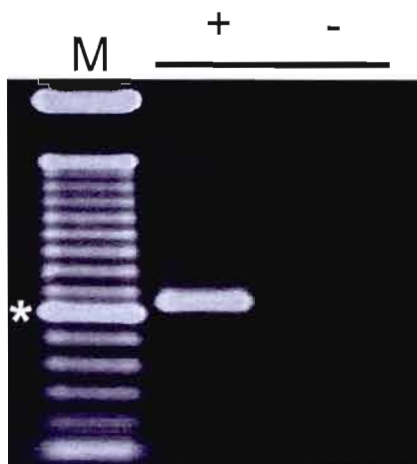


Figure 3.3

HBZ 1 MAA - - - - SGLFRCLPVSCPEDLLV - - - - - EELVDGLLSLEEELKDK 37
 APH-3 1 MARSRSGD SGYFRQMPVLAPETAIPGTVPVVYSCISTSGGSLTSRSLGPPVLTA 56
 APH-4 1 MDT - - - - REFFRGL - - - - - 10
BR2
 HBZ 38 EEEKAVLDGLLSLEEESRGRLRRGPPGEKAPPRGETHRDRQRR - - - - - EEKRK 86
 APH-3 57 EEPLTPVETLLNLCERSGRDQM - - - - DMLQGNG - KR - - PRGRPLGSSSKDLGYWK 104
 APH-4 11 - - - - - SSEDRAKKVSTPEEMPGSG - ARRARKGRPLGSGGANDGTRK 50
BR1
 HBZ 87 RKKEREKEEEKQIAEYLKRKEEEKARRRRRAEKKAADVARRKQEEQERRERKWRQG 142
 APH-3 105 SRLDDLAQEQEQLLSYWRERKKKEERRRRCREEL - - - - - EERKRQKKERRQ 150
 APH-4 51 SRLRALVDEQAEIQRLLREQEKEARRRERLQER - - - - - ERKRKEREARHQ 96
DBD
 HBZ 143 AEKAKQHSARKEKMQELGIDGYTRQLEGEVESLEAERRKLLQEKEDLMGEVNYWQG 198
 APH-3 151 QLDC - - - - - IDMLGFEGFCDLLEGYIDFLERESQQLRAGCEESLEEYFWYMD 197
 APH-4 97 QRKC - - - - - MDLLGFEGYCDLLQGYIDVLEKERKDLEEECNHNLGELFRLSA 143

 HBZ 199 RLE - - - - - AMW - - - - - LQ 206
 APH-3 198 CWGQGVDIEGDSSGLC-CTWVSSRGESHSLLLQEVADLL 235
 APH-4 144 GLEQTLDIWGDSSGLAECGWVNGRGVNHGDLLGVLAELL 182

Figure 3.4

A

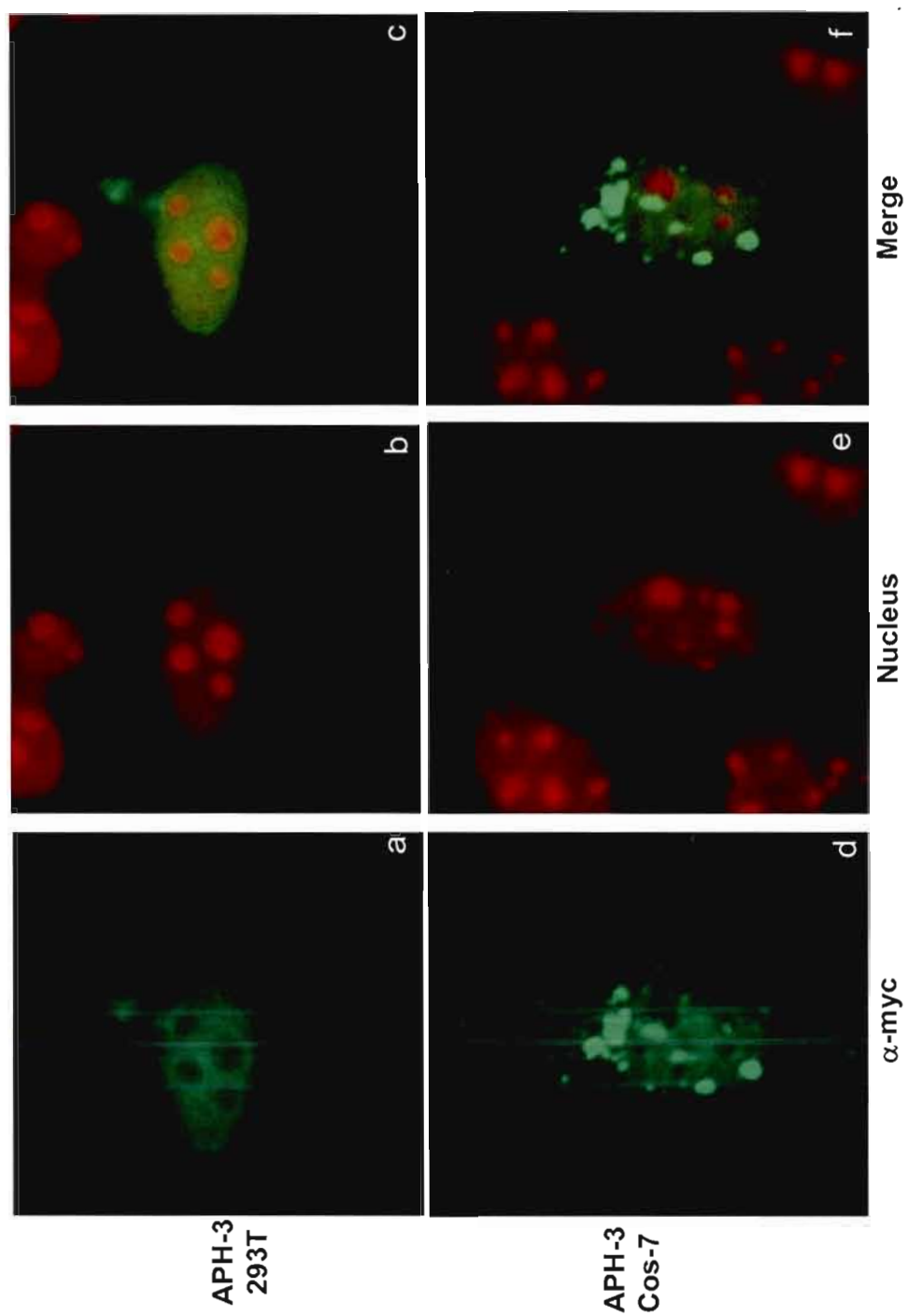


Figure 3.5

B

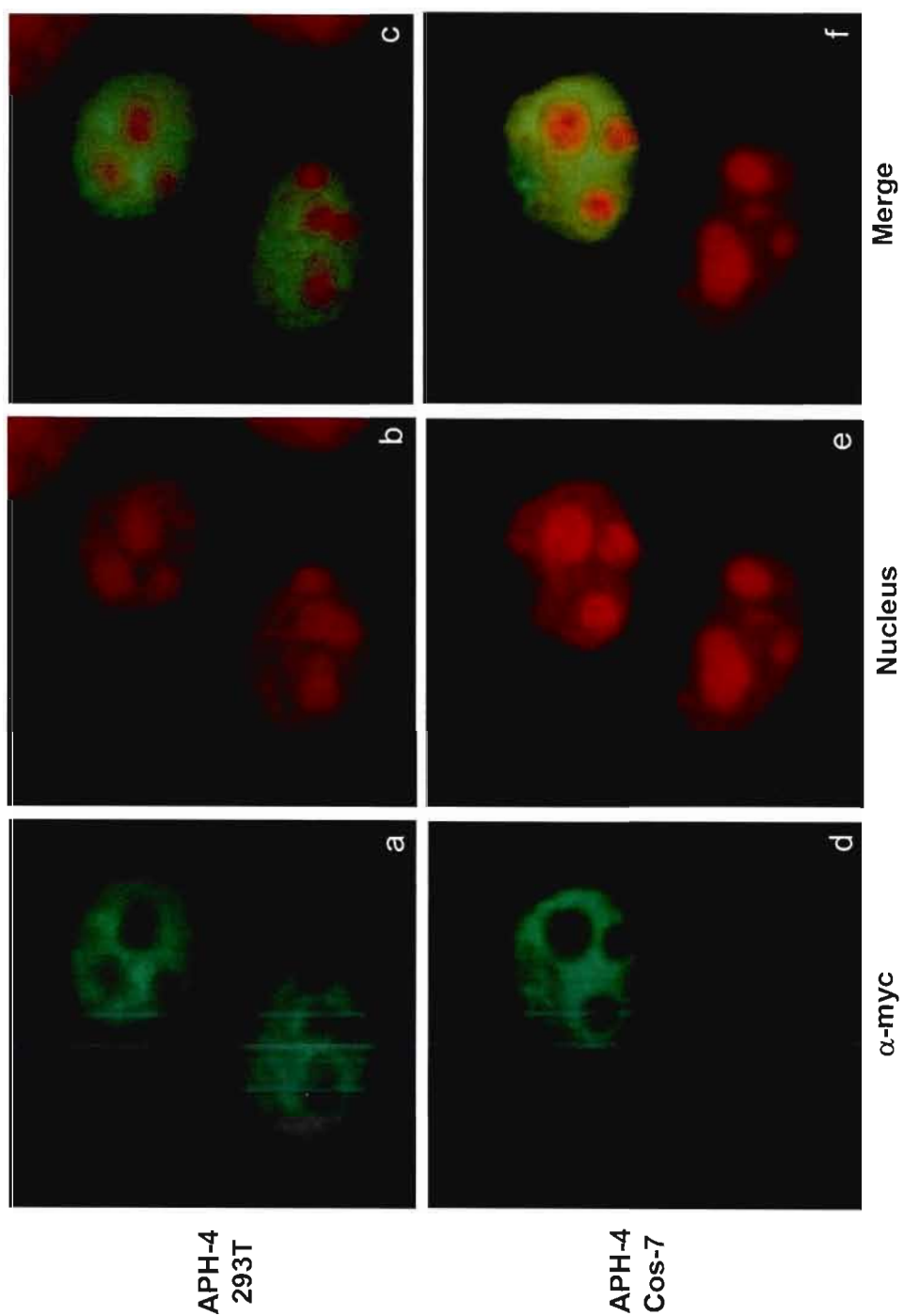


Figure 3.5

C

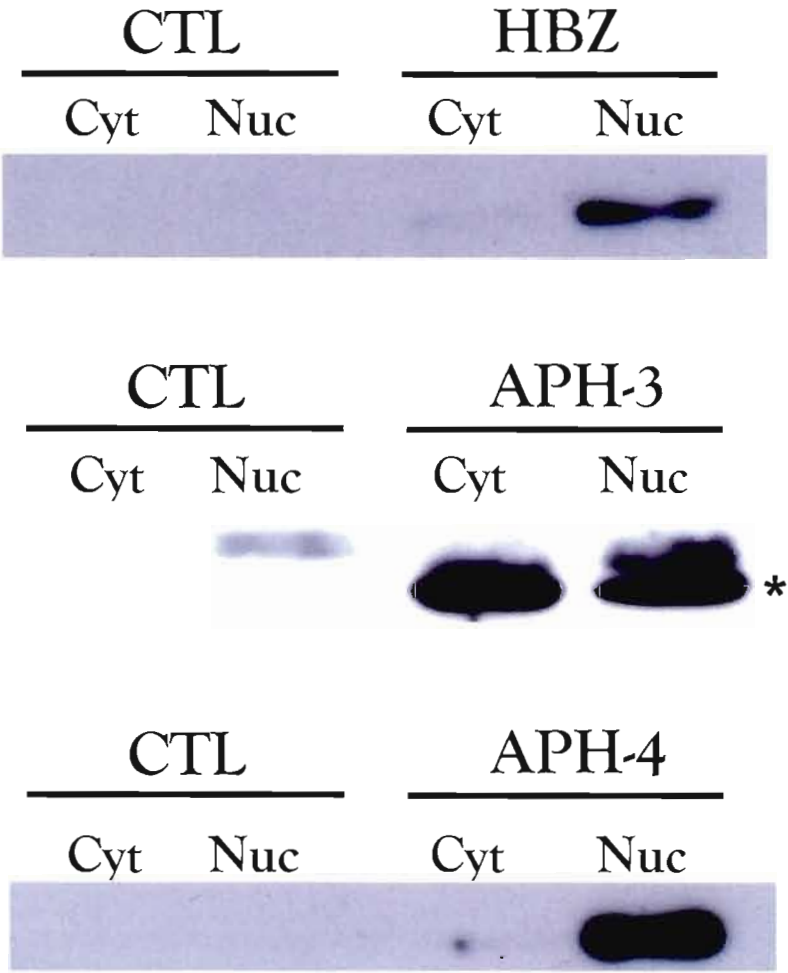


Figure 3.5

A

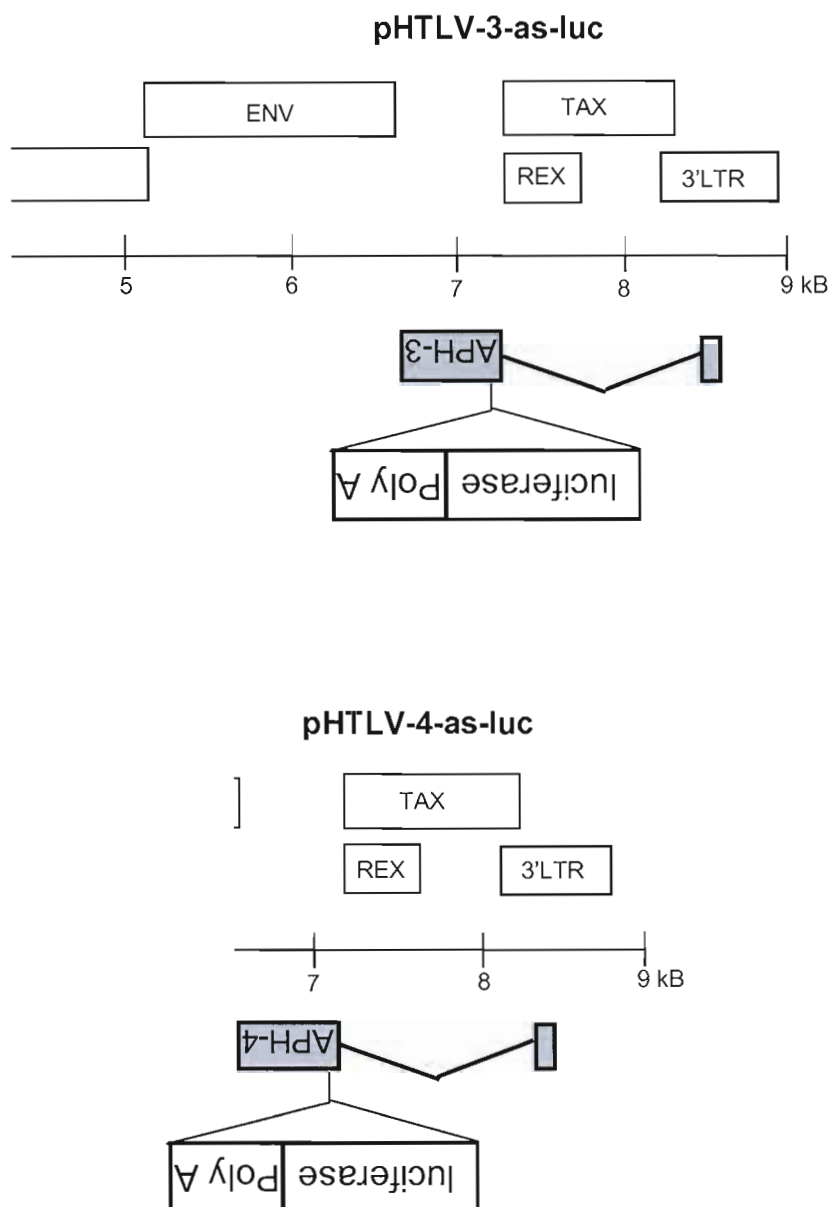


Figure 3.6

B

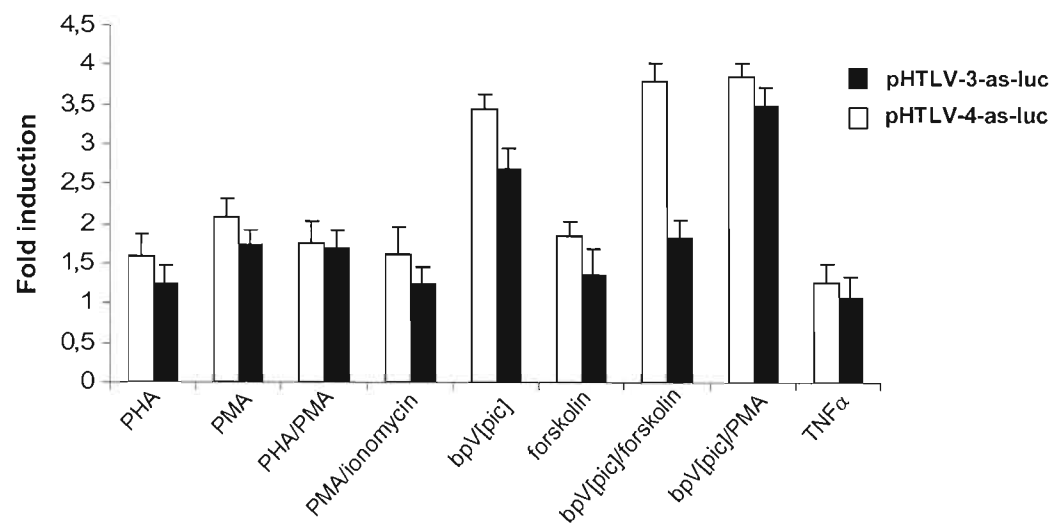


Figure 3.6

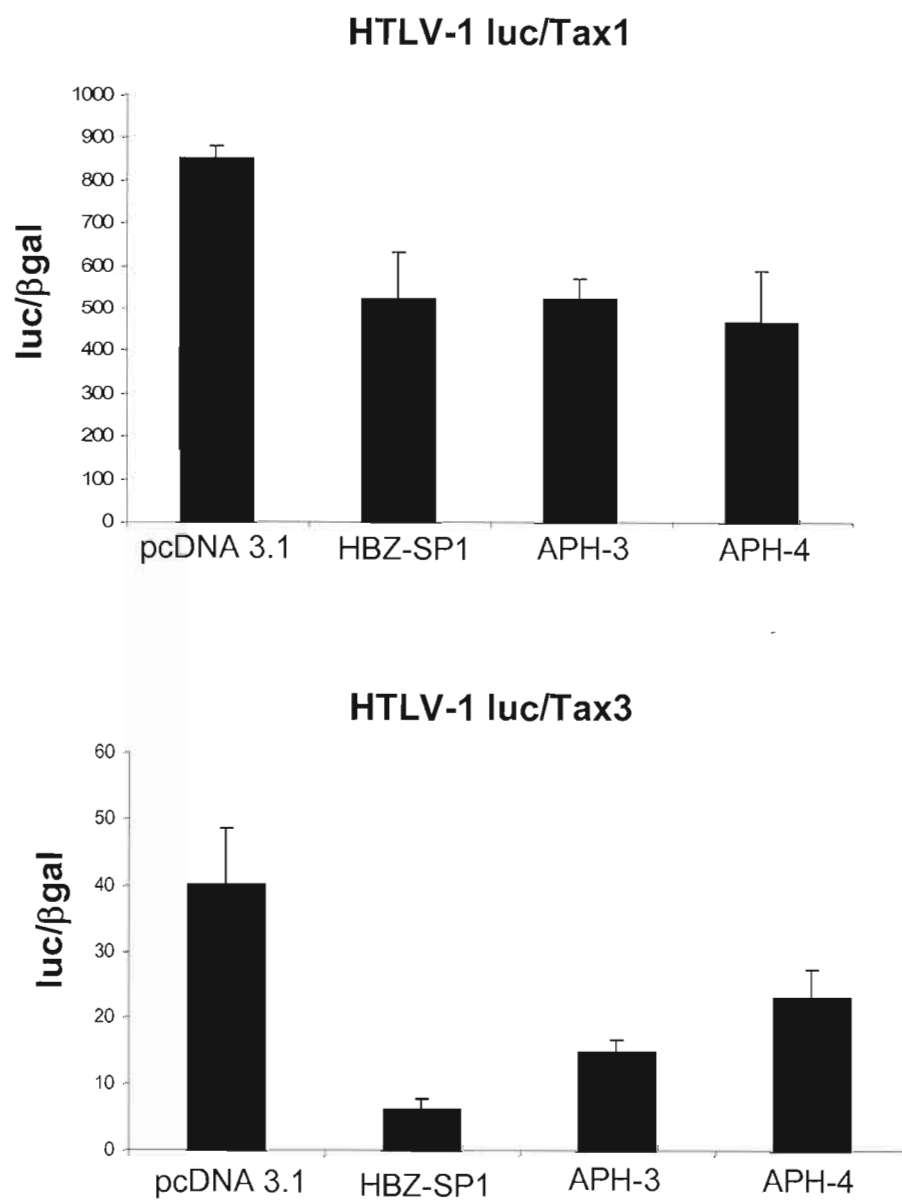


Figure 3.7

CHAPITRE IV

DISCUSSION COMPLÉMENTAIRE AUX ARTICLES ET CONCLUSIONS

Les résultats présentés dans ce mémoire mettent en lumière l'existence de transcription antisens chez les rétrovirus humains HTLV-2, -3 et -4. L'extrémité du transcrits antisens a été caractérisée chez HTLV-2, tandis que les sites de polyadénylation a été identifié pour les rétrovirus HTLV-2, -3 et -4. Les structures des transcrits semblent conservées dans la famille des rétrovirus HTLV, puisque le processus d'épissage permet l'apport d'un premier exon codant provenant du LTR3' au cadre de lecture ouvert identifié et conservé chez la grande majorité des clones dont les séquences sont disponibles sur « GenBank ». La transcription antisens chez HTLV-3 et -4 a le potentiel d'être stimulée par des activateurs des cellules T, tels que le bpV[pic] et le PMA. Des expériences de microscopie confocale dans un contexte de surexpression ont permis la détection de ces trois nouvelles protéines dans le noyau des cellules COS-7 en plus d'être également présente dans le cytoplasme en ce qui concerne APH-3. Cette localisation nucléaire s'est avérée avoir un lien avec l'une de leurs fonctions communes : l'inhibition de la transcription virale dépendante de la protéine virale Tax.

4.1 Les rétrovirus HTLV-2, -3 et -4 expriment un transcrit antisens.

Les expériences de RT-PCR permettant la détection des transcrits antisens ont fait appel à des clones proviraux complets des rétrovirus HTLV-2 (numéro d'acquisition de « GenBank » :M10060), HTLV-3 (numéro d'acquisition de « GenBank » :DQ093792) et HTLV-4 (séquence non publiée) contenus dans des plasmides. Nous avons détecté des amplicons PCR ayant des tailles plus petites que celles correspondant à des transcrits non épissés pour ces trois virus. Le séquençage de ces derniers a confirmé la provenance de ces transcrits, soient les brins antisens de ces rétrovirus. Il faut toutefois mentionner que ces

expériences ont été effectuées à partir d'un isolat de chacun de ces virus. C'est pourquoi, des analyses complémentaires telles que la détection du transcrit antisens chez des lignées cellulaires chroniquement infectées ou encore, à partir d'ARN de PBMC de patients infectés, ont été achevées dans le cas de HTLV-2. De plus, des études bioinformatiques ont permis de compléter ces observations en confirmant que les autres clones dont les séquences sont disponibles sur « GenBank » contiennent également ces cadres de lectures ouverts sur les brins complémentaires de leurs génomes. De cette manière, les onze génomes de HTLV-2 et les deux séquences provirales complètes de HTLV-3 (DQ462191 et EU649782) ont servi à faire ces analyses bioinformatiques. Il en résulte que six clones de HTLV-2 contiennent le même cadre de lecture ouvert que celui identifié lors de nos expériences (Y13051, L20734, L11456, AF326584, AF139382 et AF074965) tandis que les cinq autres (X89270, NC_001488, AF326583, AF412314 et Y14365) possèdent une version écourtée de ce dernier (constitué de respectivement 101, 135, 167, 114 et 19 acides aminés). Il est toutefois possible que ces cinq clones différents puissent contenir des erreurs de séquençage, comme cela a été le cas du clone qui a fait l'objet de la présente étude. Ensuite, la séquence du clone DQ462191 de HTLV-3 ne contient pas l'ORF antisens, puisqu'il manque 366pb dans la région pX. En raison de cette délétion, ce clone viral est également dans l'impossibilité de produire la protéine accessoire RorfII (Chevalier et al., 2007). D'un autre côté, la seconde séquence de HTLV-3 (EU649782) présente le même cadre de lecture ouvert que celui identifié. Finalement, aucune séquence supplémentaire du rétrovirus HTLV-4 n'est disponible à l'heure actuelle, excepté celle avec laquelle nous travaillons. De ce fait, toutes les données à ce jour permettent d'affirmer que ces rétrovirus ont la capacité d'exprimer un transcrit antisens. Les mêmes expériences de RT-PCR effectuées à partir de cellules transfectées par d'autres clones proviraux pourraient toutefois confirmer ces études bioinformatiques.

Les expériences de RT-PCR ont d'abord permis la détection d'un signal intense correspondant aux transcrits épissés chez les trois rétrovirus, mais à quelques reprises, l'amplification d'un très faible signal non épissé a été possible. L'obtention d'un plus fort signal non épissé n'aurait pas été surprenant, considérant que le transcrit d'HBZ présente un ratio de transcrits épissés versus non épissés d'environ 4:1 (Usui et al., 2008). Il faut toutefois préciser que les techniques utilisées lors de nos expériences de RT-PCR ne sont pas

quantitatives mais plutôt qualitatives. L'exécution de PCR en temps réel faisant appel à des amorces antisens complémentaires aux ORF identifiés, en plus d'une amorce sens situés dans le LTR3' pour l'amplification du transcrit non épissé ou une amorce chevauchant les deux exons pour l'amplification du transcrit épissé permettrait de clarifier ce point. En effet, cette sélection d'amorces permettrait d'amplifier spécifiquement l'un ou l'autre des transcrits, ce qui quantifierait l'expression de ces deux types de transcrits, s'il a lieu.

Les séquences ESE (*exonic splicing enhancer*) ainsi que les ESS (*exonic splicing silencer*) sont de petites régions situées dans les exons et elles ont la possibilité de promouvoir ou inhiber l'épissage du pré-ARN en ARNm. Considérant que les transcrits antisens de HTLV-2, -3 et -4 sont épissés suivant un unique patron, il est fort possible que ces derniers contiennent des ESE dans leurs séquences exoniques afin que la machinerie cellulaire humaine identifie adéquatement les sites d'épissages. D'un autre côté, considérant que le transcrit antisens de HTLV-1 est épissé alternativement, le mécanisme de régulation de l'épissage de chacun des transcrits doit être précisément orchestré et fait possiblement appel à des séquences ESE et ESS.

Chez HTLV-1 et VIH-1, l'exportation vers le cytoplasme de certains transcrits partiellement ou non épissés nécessite le renfort d'une autre protéine virale, soit respectivement Rex et Rev pour ces deux rétrovirus (Kashanchi and Brady, 2005). En effet, les tiges boucles formant le RexRE situées entre le signal de polyadénylation et le site de clivage des transcrits *env*, *gag*, *pol* et *pro* sont essentielles à leurs interactions avec Rex. De ce fait, il y a stabilisation de l'ARNm suivi de leur exportation vers le cytoplasme. La séquence nucléotidique à l'origine du RexRE est située entre les régions U3 et R du LTR3' de HTLV-1. La formation tridimensionnelle de ce RexRE n'est probablement pas aussi efficace sur le transcrit antisens. Conséquemment, la divergence nucléotidique du RexRE antisens comparativement à celui formé sur les transcrits sens, a probablement un impact sur une possible liaison avec la protéine Rex. Quoiqu'il en soit, il est impossible que Rex soit impliqué dans l'export des transcrits épissés puisque, chez les trois rétrovirus, l'épissage délète les transcrits antisens de la séquence nucléotidique formant le RexRE. Il est toutefois probable que les transcrits non épissés puissent bénéficier de l'aide de Rex lors de leur exportation, mais de plus amples manipulations sont nécessaires afin de s'assurer de la fonctionnalité de ce RexRE antisens.

4.2 L'initiation de la transcription a lieu dans le LTR3' et les transcrits sont polyadénylés.

Dans un premier temps, les amorces sens utilisées lors de la détection des transcrits antisens sont localisées dans le LTR3', plus précisément dans la région R. Étant donné que des signaux de RT-PCR ont été obtenus en faisant appel à ces amorces, on peut affirmer que l'initiation de la transcription a assurément lieu au moins dans cette région pour les trois rétrovirus. Afin d'obtenir plus de précision à ce sujet, des expériences de 5'RACE ont permis l'identification de plusieurs sites d'initiation de la transcription dans les régions R et U5 du transcrit antisens de HTLV-2. De tels résultats s'expliquent par l'absence de boîte TATA en amont de l'initiation de la transcription. Une analyse bioinformatique des possibles sites de liaison des facteurs transcriptionnels à l'aide du logiciel TESS (<http://www.cbil.upenn.edu/cgi-bin/teess/teess>) a prédit plusieurs sites pour différents facteurs de transcription y compris six sites Sp1. Considérant qu'il a été publié que l'activité promotrice d'HBZ est dépendante des sites de liaisons Sp1, la mutation de ces mêmes sites dans une construction contenant le promoteur de HTLV-2 en amont du gène rapporteur de la luciférase permettrait de confirmer ou d'infirmer l'importance de ces sites pour l'expression d'un transcrit antisens chez HTLV-2. Notons que ces sites probables de liaison sont différents de ceux identifiés chez HTLV-1 (Yoshida et al., 2008). Dans le même ordre d'idée, des expériences de 5'RACE sont présentement en cours dans notre laboratoire afin d'identifier les sites d'initiation de la transcription des transcrits antisens de HTLV-3 et -4. Quelques sites ont déjà été identifiés pour HTLV-3, soit aux positions -30, -37 et -198 (numérotation débutant au nucléotide A de l'ATG initiateur de la traduction), mais des séquences supplémentaires sont nécessaires à une meilleure caractérisation. Ce type de manipulation est complexe puisqu'on estime que la transcription antisens est beaucoup moins abondante que la transcription sens. Par exemple, chez HTLV-1, on estime qu'il y a entre cinquante et cent fois moins de transcrits antisens que de transcrits codant pour la protéine Tax (Larocca et al., 1989). De plus, l'amplification de l'extrémité 5' d'un transcrit est beaucoup plus fastidieuse que l'identification d'un site de polyadénylation. En effet, les trousseaux commerciaux permettant l'amplification de l'extrémité 3' d'un transcrit proposent simplement une réaction de transcription inverse faisant appel à un oligo (dT) comprenant une séquence adaptatrice à

son extrémité 5' qui est utilisée lors de la PCR nichée. D'un autre côté, les méthodes d'amplification des extrémités 5' sont plus complexes. Par exemple, la trousse RLM-RACE (Ambion) propose de faire subir à l'échantillon d'ARN une série de réactions enzymatiques permettant le clivage des extrémités 5'-phosphate et «cap» en plus d'effectuer une extraction au phénol/chloroforme et une ligation d'un adaptateur avant la réaction de RT. De plus, le matériel utilisé pour ces expériences est issu d'une extraction d'ARN de cellules transfectées par un plasmide contenant l'ADN proviral des rétrovirus et plusieurs traitements DNase sont nécessaires à l'obtention d'ARN exempt de plasmide. Néanmoins, étant donné qu'aucune boîte TATA n'a été repérée en amont des ORF antisens de HTLV-3 et -4, le même patron d'initiation de la transcription devrait être observé. Les analyses bioinformatiques des promoteurs de ces derniers par le logiciel TESS a mis en lumière deux sites Sp1 possibles pour HTLV-3 et quatre pour HTLV-4. Notons que ces derniers sont différents de ceux prédits pour HTLV-1 et -2. Une étude exhaustive de l'activité promotrice sera essentielle à la compréhension du patron d'expression de ces transcrits antisens.

Ces recherches ont également permis l'identification d'un site fonctionnel de polyadénylation pour les transcrits antisens des rétrovirus HTLV-2, -3 et -4. Dans ces trois cas, le signal de polyadénylation est consensus (AAUAAA) et une région riche en nucléotides G et U se retrouve en aval du site de clivage. Des comparaisons entre les séquences des clones disponibles sur « GenBank » permettent de confirmer que la totalité des clones précédemment analysés contiennent les mêmes signaux de polyadénylation, sites de clivages et régions riches en G et U que ceux identifiés par les expériences de 3'RACE, excepté un clone d'HTLV-2 (AF139382) qui ne possède pas un signal consensus de polyadénylation (AAUAGA). Considérant que dix génomes d'HTLV-2 sur onze contiennent ce signal de polyadénylation consensus, l'hypothèse d'une erreur de séquençage est possible pour le clone divergent. De plus, notons que ce dernier contient l'ORF originalement identifié et que ses sites donneur et accepteur d'épissage sont conservés.

4.3 L'épissage des transcrits engendre des cadres de lecture ouverts distribués sur deux exons.

Le séquençage des amplifications PCR lors de la détection des transcrits antisens de HTLV-2, -3 et -4 ont permis de constater que ces derniers sont épissés. En effet, le patron d'épissage s'apparente à celui du transcrit d'HBZ SP1 (Cavanagh et al., 2006). Pour les rétrovirus HTLV-1, -2 et -4, les premiers exons qui sont localisés dans la région R du LTR3', fournissent treize nucléotides, comparativement à vingt-huit chez HTLV-3. Ainsi, il y a production de protéines qui ont été nommées APH-2, -3 et -4 pour « *antisense protein of HTLV* ». Ces dernières sont constituées de respectivement 183, 235 et 182 acides aminés. Des analyses de séquences ont également été effectuées sur les différents clones disponibles et les sites donneurs et accepteurs d'épissage sont conservés chez la totalité des clones disponibles, excepté chez deux isolats d'HTLV-2. Il s'agit du clone AF074965 qui a une mutation du site donneur d'épissage (ATGT au lieu de AGGT) et le clone Y14365 dont son site accepteur d'épissage est différent (AGGA comparativement à AGAC chez les autres clones). La mutation du clone Y14365 engendre probablement un épissage correct puisque la littérature propose un site accepteur d'épissage consensus constitué des nucléotides AGG (Alberts et al., 2002). D'un autre côté, la mutation du site donneur d'épissage du clone AF074965 n'empêche probablement pas l'épissage désiré, mais de plus amples études sont nécessaires avant de conclure sur ce sujet.

4.4 APH-2, -3 et -4 se localisent dans le noyau en plus d'être présents dans le cytoplasme pour APH-3.

Les protéines résultant des cadres de lectures ouverts contenus dans chacun des transcrits antisens des rétrovirus HTLV-2, -3 et -4 ont fait l'objet d'expériences de localisation cellulaire. Dans un premier temps, la protéine APH-2, qui a été étiquetée en C-terminal par une étiquette Myc ou GFP, a été transfectée dans des cellules COS-7 ou HeLa. Les résultats présentés dans le premier article font état d'une localisation principalement nucléaire sans co-localisation avec les nucléoles, contrairement à HBZ-SP1, qui présente une distribution plus prononcée dans les nucléoles (Murata et al., 2006). Un fort signal est

également présent à proximité du noyau sous forme d'un point intense. Cette localisation extranucléaire pourrait correspondre à une vacuole de dégradation, compte tenu que ces expériences ont été effectuées dans un contexte de surexpression de la protéine APH-2. Des manipulations supplémentaires de co-localisation avec une protéine formant le protéasome tel que le protéasome 20S permettraient la spécification de la nature de ce signal. De plus, l'ajout d'un composé inhibant la dégradation par le protéasome tel que le MG132 aux cellules en cultures, avant l'observation en microscopie, pourrait également confirmer ou infirmer cette hypothèse. L'expression de APH-2 a été authentifiée par Western blot et cette dernière migre à une taille approximative de 25 kDa, comparativement à un poids moléculaire attendu de 21,02 kDa. Ensuite, la protéine APH-3 a été fusionnée à une étiquette Myc en *N*- et en *C*-terminal. Dans ces deux cas, elle a été détectée dans le noyau et dans le cytoplasme des cellules 293T et COS-7. Le fait que le patron de détection de APH-3 en microscopie confocale ne diffère pas lorsque l'étiquette est située à l'une ou l'autre des extrémités, écarte l'hypothèse que cette dernière soit clivée. Finalement, la protéine APH-4 a été étiquetée de la même manière que APH-3 et sa distribution s'est avérée strictement nucléaire dans les cellules 293T et COS-7, sans toutefois co-localiser avec les nucléoles. L'expression et la distribution de APH-3 et -4 ont été confirmées par Western blot à partir d'extraits cellulaires fractionnés. En effet, la migration des fractions cytoplasmiques et nucléaires des cellules transfectées par l'un ou l'autre des vecteurs d'expression de APH-3 ou -4 ont confirmé la répartition de ces protéines, soit nucléaire et cytoplasmique dans le cas d'APH-3 et strictement nucléaire pour APH-4. Ces dernières ont migré au poids moléculaire apparent de 29 kDa pour la protéine APH-3 et de 17,5 kDa pour la protéine APH-4, comparativement à des tailles attendues de respectivement 28,23 kDa et 22,17 kDa. Il faut toutefois noter qu'une bande supplémentaire migrant au dessus de APH-3 et qui est présente dans les pistes contenant les extraits cellulaires de cellules non transfectées semble s'intensifier dans les extraits de cellules transfectées par le vecteur d'expression de APH-3. Les résultats obtenus jusqu'à présent ne permettent pas de confirmer la spécificité de cette bande supplémentaire. Il est toutefois possible d'émettre l'hypothèse que APH-3 subisse une modification post-traductionnelle, ce qui expliquerait que cette forme d'APH-3 ait un poids moléculaire légèrement plus élevé que APH-3 originale. Le fait que cette bande supplémentaire migre à la même hauteur qu'un signal non spécifique rend l'interprétation ardue. Une autre hypothèse

pourrait expliquer la présence de cette bande supplémentaire : APH-3 engendrerait des modifications dans les cellules transfectées, par exemple en se liant à des protéines cellulaires ou en activant la transcription de certains gènes, ce qui augmenterait leur concentration cellulaire. En conséquence, s'il y a surproduction d'une certaine protéine et que cette dernière est reconnue par l'anticorps employé lors du Western blot, il se peut qu'une bande supplémentaire apparaisse. Des expériences complémentaires consistant en un immunobuvardage de type Western blot à partir de cellules transfectées par le vecteur d'expression APH-3 étiqueté différemment (par exemple, avec une étiquette GFP) permettrait l'élimination du signal non spécifique engendré par l'utilisation de l'anticorps précédent.

Il a été démontré que le signal de localisation nucléaire d'HBZ est dû à la présence de trois motifs distincts, soit le domaine de liaison à l'ADN ainsi que des deux régions basiques (BR1 et BR2) (Hivin et al., 2005). Des études bioinformatiques ont identifié des régions constituées d'acides aminés basiques (arginine, histidine et lysine), plus précisément, les résidus 50 à 96, 122 à 149 et 50 à 99 dans respectivement les protéines APH-2, -3 et -4. Ces régions pourraient être à l'origine de cette localisation nucléaire. La production de mutants de délétions de ces régions apparemment basiques suivie de leur observation en microscopie confocale confirmerait peut-être cette hypothèse.

Les observations en microscopie confocale ont été effectuées dans deux types cellulaires différents, afin de discriminer l'obtention d'un signal non spécifique caractéristique à un type cellulaire. Ces expériences ont également été réalisées dans une lignée lymphocytaire humaine (Jurkat E6.1), qui constitue le type cellulaire infecté par ces rétrovirus, mais aucune conclusion n'a pu être prise. En effet, les cellules Jurkat E6.1, comme la totalité des cellules lymphocytaires ont une morphologie particulière caractérisée par un noyau prédominant rendant difficile la distinction du cytoplasme. Un signal a néanmoins été obtenu dans cette lignée cellulaire, ce qui confirme que les protéines APH-2, -3 et -4 peuvent s'exprimer dans des lymphocytes.

La génération de constructions constituées de l'ADN proviral des rétrovirus HTLV-2, -3 et -4 auxquels serait ajouté une étiquette Myc en C-terminal des ORF antisens permettrait d'exprimer ces protéines dans un contexte s'apparentant plus à une infection rétrovirale. Ainsi, la production de ces protéines dépendrait de l'activité promotrice des LTR3' de ces rétrovirus et le ratio entre la production des différentes protéines virales serait respecté. De ce

fait, les études de localisation cellulaire ainsi que la reprise des immunobuvardages de type Western blot à partir des cellules transfectées par ces constructions confirmeraient les résultats précédemment obtenus.

4.5 APH-2, -3 et -4 inhibent la transcription de HTLV-1 dépendante de la protéine virale Tax.

Il a précédemment été mis en lumière que la protéine antisens de HTLV-1 a le potentiel d'inhiber la transcription virale dépendante de Tax, par l'intermédiaire de son domaine *leucine zipper* ainsi que par deux régions composées des acides aminés LXXLL. Les résultats obtenus démontrent que la protéine APH-2 a la capacité d'inhiber la transcription de HTLV-2 dépendante de Tax2. Il semble que ce blocage est plus important lorsqu'il y a une quantité croissante d'APH-2. De plus, le second article fait état du potentiel inhibiteur de APH-3 et -4 sur la transcription de HTLV-1 dépendante des protéines Tax 1 et Tax3. Ces expériences ont été réalisées puisque Tax1 et Tax3 ont précédemment été reconnues pour activer la transcription de HTLV-1 et -3 (Bex and Gaynor, 1998; Calattini et al., 2006). Également, puisqu'il est convenu que Tax3 active la transcription de HTLV-3, il serait intéressant de confirmer le potentiel inhibiteur d'APH-3 sur la transcription d'HTLV-3 dépendante de Tax3. Il serait aussi pertinent de vérifier si Tax4 est capable d'activer la transcription du LTR5' de HTLV-4 et si tel est le cas, d'observer si APH-4 inhibe également cette transcription. À la lumière de ces résultats, il est possible d'affirmer que ces trois nouvelles protéines antisens ont la capacité d'inhiber la transcription virale dépendante de Tax.

4.6 Le motif leucine zipper ne semble pas présent chez APH-2, -3 et -4.

Considérant l'effet inhibiteur de la transcription virale dépendante de Tax de APH-2, -3 et -4, des études bioinformatiques ont été réalisées afin d'identifier la ou les régions possiblement impliquées dans cette inhibition. De la sorte, aucune prédiction de motif *leucine zipper* n'a été possible à partir du site Internet <http://2zip.molgen.mpg.de/>, comme c'est le cas

pour HBZ. En plus de vérifier la présence des résidus leucines, ce site permet la prédiction de la formation d'hélice α , également nécessaire à la formation d'un *leucine zipper*. Par contre certains résidus leucines séparés par six, sept ou huit acides aminés semblent conservés chez ces trois protéines, pouvant laisser présager l'existence d'une version atypique d'un *leucine zipper*. La mutation des leucines impliquées dans le potentiel *leucine zipper* des protéines APH-2, -3 et -4 est essentielle à la compréhension des mécanismes d'interactions permettant l'inhibition de la transcription virale. Ainsi, la reprise des expériences visant à mettre en évidence le potentiel d'inhibition de la transcription virale dépendante de Tax à partir des constructions APH mutées permettrait de s'assurer que cet alignement de leucines n'est pas impliqué dans un quelconque *leucine zipper*. De plus, ces vecteurs d'expression contenant les ORF antisens des rétrovirus HTLV-2, -3 et -4 mutés ou pas, pourraient être utilisés lors d'expériences de retard sur gel ou d'immunoprécipitation avec des facteurs de transcription cellulaires connus comme interagissant avec HBZ, tels que JunB, c-Jun, JunD, Fos, CREB et CREB-2.

Il a été démontré que les domaines LXXLL d'HBZ peuvent également engendrer une inhibition de la transcription virale dépendante de Tax *via* une interaction directe avec CBP/p300 (Clerc et al., 2008). Considérant qu'il n'y a aucune prédiction de *leucine zipper* chez les protéines APH-2, -3 et -4 et qu'ils ont tout de même l'aptitude de réguler la transcription, la présence de tels motifs pourrait expliquer ces résultats. Il s'avère que APH-2 contient un motif LXXLL et un autre pouvant s'y apparenter (IXXLL), APH-3 renferme deux domaines VXXLL, tandis que APH-4 comporte un motif LXXLL et un IXXLL. Ces domaines particuliers pourraient potentiellement être impliqués dans des interactions avec CBP/p300 et ainsi être responsables de la fonctionnalité de ces protéines. De plus amples manipulations en laboratoire devraient permettre d'authentifier ou non cette hypothèse. Ainsi, la mutation de chacun des ces domaines LXXLL dans les vecteurs d'expressions APH-2, -3 et -4 suivi de leur transfection en présence des différents Tax et pHTLV-luc permettraient de mettre la lumière sur l'effet inhibiteur de ces nouvelles protéines antisens.

4.7 La transcription antisens est modulable par des agents activateurs des cellules T.

L'étude de l'activité promotrice antisens des rétrovirus HTLV-3 et -4 a nécessité la génération de constructions contenant le gène de la luciférase inséré dans le deuxième exon des ORF de APH-3 et -4. Étant donné qu'il a préalablement été mis en évidence que la transcription antisens de VIH-1 est modulée positivement par des agents activateurs des cellules T, la transfection de ces vecteurs dans des cellules lymphocytaires Jurkat E6.1 suivie d'une stimulation par une série de composés a démontré des résultats comparables (Landry et al., 2007). Ainsi, les transcriptions antisens de HTLV-3 et -4 sont stimulés significativement par un agent activateur des protéines kinases, le PMA, ainsi que par un inhibiteur des protéines phosphotyrosines phosphatases, le bpV[pic] ou une combinaison de ces deux produits. De plus, le LTR3' de HTLV-3 est également activé par la forskoline, qui est un composé qui augmente la concentration cellulaire d'AMP cyclique, seule ou en présence de bpV[pic].

D'autre part, ces constructions ont été transfectées dans des cellules 293T en présence des protéines Tax3 et Tax4 et aucune augmentation significative de la transcription antisens n'a été observée chez HTLV-3 et -4 (données non présentées). Contrairement à ce qui a été établi concernant la transcription antisens de HTLV-1, les promoteurs de APH-3 et -4 ne semblent pas être sujet à une trans-activation par leurs protéines Tax respectives (Yoshida et al., 2008). Il serait toutefois important de vérifier si le site d'intégration influence ces résultats, comme ce fut démontré pour HTLV-1 (Landry et al., 2008), en transfectant Tax3 et Tax4 dans des cellules exprimant stablement les constructions pHTLV-3-as-luc et pHTLV-4-as-luc. La trans-activation de la transcription dépendante de Tax nécessite la formation d'un dimère entre Tax et un facteur de transcription cellulaire, qui se lie aux séquences CRE situées dans les TRE du LTR viral. Puisque les deux LTR sont identiques, les TRE sont également présents dans le LTR3'. De ce fait, Tax aurait potentiellement la possibilité de jouer un rôle dans la transcription antisens. Par contre, les TRE sont situés dans la région U3 et cette dernière est en aval des sites d'initiation de la transcription des trois rétrovirus étudiés. C'est également le cas chez HTLV-1 et il a tout de même été démontré que Tax active la transcription antisens. Il est donc fort probable que l'activation de la

transcription soit entre autre dépendante d'éléments de régulation situés en aval des sites d'initiation de la transcription. La modulation de la transcription par Tax pourrait néanmoins diverger entre ces rétrovirus, puisque le LTR de HTLV-3 ne contient que deux TRE contrairement aux LTR de HTLV-1 et -2 qui en contiennent trois (Calattini et al., 2006).

Une étude plus approfondie de l'activité promotrice antisens pourra être faite à partir des constructions luciférase préalablement décrites. Par exemple, la mesure de l'activité luciférase suite à des délétions progressives des LTR3' ciblerait les régions cruciales des promoteurs de ces derniers. Conséquemment, les facteurs transcriptionnels impliqués dans l'expression de ces protéines pourraient être identifiés.

4.8 Conclusion

Les quatre rétrovirus HTLV infectent les lymphocytes T mais seul HTLV-1 est reconnu comme induisant des ATL tandis que HTLV-1 et -2 induisent des désordres neurologiques. De récentes études ont démontré qu'il y a une sélection préférentielle des intégrations rétrovirales de HTLV-1 contenant des LTR5' hyperméthylés ou absents lors de l'oncogenèse, ce qui implique que la totalité des protéines virales ne s'expriment plus. Par contre, le LTR3' semble constamment conservé, ce qui permet l'expression de la protéine antisens HBZ. Ainsi, ces faits laissent présager que cette dernière pourrait potentiellement jouer un rôle dans le processus d'initiation ou de persistance de l'oncogenèse. Considérant que les rétrovirus HTLV-2, -3 et -4 ne sont pas associés au développement de quelconque pathogénicité, l'étude de la transcription antisens de ces derniers a été essentielle. Ainsi, ces résultats permettent d'affirmer que les rétrovirus HTLV-2, -3 et -4 expriment également un transcrit antisens qui présente le même patron d'épissage que HBZ, sans toutefois être alternativement épissé. Ces protéines se localisent entre autres dans le noyau, sans toutefois être distribuées de la même manière que HBZ, ce qui implique que des interactions avec différentes protéines cellulaires sont probables. D'un autre côté, malgré le fait que les protéines codées par ces transcrits ont la capacité d'inhiber la transcription virale dépendante de Tax, elles sont apparemment différentes d'HBZ puisqu'elles ne semblent pas contenir de *leucine zipper*. Les régions LXXLL contenues en deux exemplaires dans chacune de ces protéines antisens sont potentiellement responsables de leurs rôles de régulation de la

transcription virale. De plus amples études menées sur les régions fonctionnelles et structurales de ces protéines sont indispensables à la compréhension des mécanismes d'infection, de réplication virale et de pathogénèse de ces rétrovirus.

RÉFÉRENCES

- Akoulitchev, S., Makela, T.P., Weinberg, R.A., and Reinberg, D. (1995). Requirement for TFIIF kinase activity in transcription by RNA polymerase II. *Nature* 377, 557-560.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002). *Molecular Biology of the Cell*. New York and London, Garland Science.
- Arisawa, K., Soda, M., Endo, S., Kurokawa, K., Katamine, S., Shimokawa, I., Koba, T., Takahashi, T., Saito, H., Doi, H., *et al.* (2000). Evaluation of adult T-cell leukemia/lymphoma incidence and its impact on non-Hodgkin lymphoma incidence in southwestern Japan. *Int J Cancer* 85, 319-324.
- Arnold, J., Yamamoto, B., Li, M., Phipps, A.J., Younis, I., Lairmore, M.D., and Green, P.L. (2006). Enhancement of infectivity and persistence in vivo by HBZ, a natural antisense coded protein of HTLV-1. *Blood* 107, 3976-3982.
- Arnold, J., Zimmerman, B., Li, M., Lairmore, M.D., and Green, P.L. (2008). Human T-cell Leukemia Virus Type-1 Antisense-encoded Gene, Hbz, Promotes T Lymphocyte Proliferation. *Blood*.
- Barbeau, B., and Mesnard, J.M. (2007). Does the HBZ gene represent a new potential target for the treatment of adult T-cell leukemia? *Int Rev Immunol* 26, 283-304.
- Bartoe, J.T., Albrecht, B., Collins, N.D., Robek, M.D., Ratner, L., Green, P.L., and Lairmore, M.D. (2000). Functional role of pX open reading frame II of human T-lymphotropic virus type 1 in maintenance of viral loads in vivo. *J Virol* 74, 1094-1100.
- Basbous, J., Arpin, C., Gaudray, G., Piechaczyk, M., Devaux, C., and Mesnard, J.M. (2003). The HBZ factor of human T-cell leukemia virus type I dimerizes with transcription factors JunB and c-Jun and modulates their transcriptional activity. *J Biol Chem* 278, 43620-43627.
- Bex, F., and Gaynor, R.B. (1998). Regulation of gene expression by HTLV-I Tax protein. *Methods* 16, 83-94.
- Bialek, R., Cirera, A.C., Herrmann, T., Aepinus, C., Shearn-Bochsler, V.I., and Legendre, A.M. (2003). Nested PCR assays for detection of *Blastomyces dermatitidis* DNA in paraffin-embedded canine tissue. *J Clin Microbiol* 41, 205-208.
- Böhnlein, S., Hauber, J., and Cullen, B.R. (1989). Identification of a U5-specific sequence required for efficient polyadenylation within the human immunodeficiency virus long terminal repeat. *J Virol* 63, 421-424.

- Bouziane, M., Cherny, D.I., Mouscadet, J.-F., and Auclair, C. (1996). Alternate strand DNA triple helix-mediated inhibition of HIV-1 U5 long terminal repeat integration *in vitro*. *J Biol Chem* 271, 10359-10364.
- Briquet, S., Richardson, J., Vanhee-Brossollet, C., and Vaquero, C. (2001). Natural antisense transcripts are detected in different cell lines and tissues of cats infected with feline immunodeficiency virus. *Gene* 267, 157-164.
- Briquet, S., and Vaquero, C. (2002). Immunolocalization studies of an antisense protein in HIV-1-infected cells and viral particles. *Virology* 292, 177-184.
- Bukrinsky, M.I., and Etkin, A.F. (1990). Plus strand of the HIV provirus DNA is expressed at early stages of infection. *AIDS Res Hum Retroviruses* 6, 425-426.
- Bunick, D., Zandomeni, R., Ackerman, S., and Weinmann, R. (1982). Mechanism of RNA polymerase II-specific initiation of transcription in vitro: ATP requirement and uncapped runoff transcripts. *Cell* 29, 877-886.
- Buratowski, S., Sopta, M., Greenblatt, J., and Sharp, P.A. (1991). RNA polymerase II-associated proteins are required for a DNA conformation change in the transcription initiation complex. *Proc Natl Acad Sci U S A* 88, 7509-7513.
- Burmeister, T. (2001). Oncogenic retroviruses in animals and humans. *Rev Med Virol* 11, 369-380.
- Calattini, S., Chevalier, S.A., Duprez, R., Afonso, P., Froment, A., Gessain, A., and Mahieux, R. (2006). Human T-cell lymphotropic virus type 3: complete nucleotide sequence and characterization of the human tax3 protein. *J Virol* 80, 9876-9888.
- Calattini, S., Chevalier, S.A., Duprez, R., Bassot, S., Froment, A., Mahieux, R., and Gessain, A. (2005). Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa. *Retrovirology* 2, 30.
- Cavanagh, M.H., Landry, S., Audet, B., Arpin-Andre, C., Hivin, P., Pare, M.E., Thete, J., Wattel, E., Marriott, S.J., Mesnard, J.M., *et al.* (2006). HTLV-I antisense transcripts initiating in the 3'LTR are alternatively spliced and polyadenylated. *Retrovirology* 3, 15.
- Chen, I.S., Quan, S.G., and Golde, D.W. (1983). Human T-cell leukemia virus type II transforms normal human lymphocytes. *Proc Natl Acad Sci U S A* 80, 7006-7009.
- Chevalier, S.A., Ko, N.L., Calattini, S., Mallet, A., Prevost, M.C., Kehn, K., Brady, J.N., Kashanchi, F., Gessain, A., and Mahieux, R. (2008). Construction and characterization of a human T-cell lymphotropic virus type 3 infectious molecular clone. *J Virol* 82, 6747-6752.

- Chevalier, S.A., Meertens, L., Calattini, S., Gessain, A., Kiemer, L., and Mahieux, R. (2005). Presence of a functional but dispensable nuclear export signal in the HTLV-2 Tax protein. *Retrovirology* 2, 70.
- Chevalier, S.A., Meertens, L., Pise-Masison, C., Calattini, S., Park, H., Alhaj, A.A., Zhou, M., Gessain, A., Kashanchi, F., Brady, J.N., *et al.* (2006). The tax protein from the primate T-cell lymphotropic virus type 3 is expressed in vivo and is functionally related to HTLV-1 Tax rather than HTLV-2 Tax. *Oncogene* 25, 4470-4482.
- Chevalier, S.A., Walic, M., Calattini, S., Mallet, A., Prevost, M.C., Gessain, A., and Mahieux, R. (2007). Construction and characterization of a full-length infectious simian T-cell lymphotropic virus type 3 molecular clone. *J Virol* 81, 6276-6285.
- Clerc, I., Polakowski, N., Andre-Arpin, C., Cook, P., Barbeau, B., Mesnard, J.M., and Lemasson, I. (2008). An interaction between the human T cell leukemia virus type 1 basic leucine zipper factor (HBZ) and the KIX domain of p300/CBP contributes to the down-regulation of tax-dependent viral transcription by HBZ. *J Biol Chem* 283, 23903-23913.
- Cockerell, G.L., Rovnak, J., Green, P.L., and Chen, I.S. (1996). A deletion in the proximal untranslated pX region of human T-cell leukemia virus type II decreases viral replication but not infectivity in vivo. *Blood* 87, 1030-1035.
- Coffin, J.M., Hughes, S.H. and Varmus, H.E. (1997). *Retroviruses*. Plainview (NY), Cold Spring Harbor Laboratory Press.
- Collins, N.D., Newbound, G.C., Albrecht, B., Beard, J.L., Ratner, L., and Lairmore, M.D. (1998). Selective ablation of human T-cell lymphotropic virus type 1 p12I reduces viral infectivity in vivo. *Blood* 91, 4701-4707.
- Conaway, J.W., and Conaway, R.C. (1991). Initiation of eukaryotic messenger RNA synthesis. *J Biol Chem* 266, 17721-17724.
- Corcelette, S., Masse, T., and Madjar, J.J. (2000). Initiation of translation by non-AUG codons in human T-cell lymphotropic virus type I mRNA encoding both Rex and Tax regulatory proteins. *Nucleic Acids Res* 28, 1625-1634.
- Denicourt, C., Edouard, E., and Rassart, E. (1999). Oncogene activation in myeloid leukemias by Graffi murine leukemia virus proviral integration. *J Virol* 73, 4439-4442.
- Denicourt, C., Kozak, C.A., and Rassart, E. (2003). Gris1, a new common integration site in Graffi murine leukemia virus-induced leukemias: overexpression of a truncated cyclin D2 due to alternative splicing. *J Virol* 77, 37-44.
- Denicourt, C., Legault, P., McNabb, F.A., and Rassart, E. (2008). Human and mouse cyclin D2 splice variants: transforming activity and subcellular localization. *Oncogene* 27, 1253-1262.

Derse, D., Hill, S.A., Lloyd, P.A., Chung, H., and Morse, B.A. (2001). Examining human T-lymphotropic virus type 1 infection and replication by cell-free infection with recombinant virus vectors. *J Virol* 75, 8461-8468.

Diop, S., Calattini, S., Abah-Dakou, J., Thiam, D., Diakhate, L., and Gessain, A. (2006). Seroprevalence and molecular epidemiology of human T-Cell leukemia virus type 1 (HTLV-1) and HTLV-2 in blood donors from Dakar, Senegal. *J Clin Microbiol* 44, 1550-1554.

Fan, N., Gavalchin, J., Paul, B., Wells, K.H., Lane, M.J., and Poiesz, B.J. (1992). Infection of peripheral blood mononuclear cells and cell lines by cell-free human T-cell lymphoma/leukemia virus type I. *J Clin Microbiol* 30, 905-910.

Feuer, G., and Green, P.L. (2005). Comparative biology of human T-cell lymphotropic virus type 1 (HTLV-1) and HTLV-2. *Oncogene* 24, 5996-6004.

Fortin, J.F., Cantin, R., Lamontagne, G., and Tremblay, M. (1997). Host-derived ICAM-1 glycoproteins incorporated on human immunodeficiency virus type 1 are biologically active and enhance viral infectivity. *J Virol* 71, 3588-3596.

Fujii, M., Niki, T., Mori, T., Matsuda, T., Matsui, M., Nomura, N., and Seiki, M. (1991). HTLV-1 Tax induces expression of various immediate early serum responsive genes. *Oncogene* 6, 1023-1029.

Gabet, A.S., Kazanji, M., Couppie, P., Clity, E., Pouliquen, J.F., Sainte-Marie, D., Aznar, C., and Wattel, E. (2003). Adult T-cell leukaemia/lymphoma-like human T-cell leukaemia virus-1 replication in infective dermatitis. *Br J Haematol* 123, 406-412.

Gabet, A.S., Mortreux, F., Charneau, P., Riou, P., Duc-Dodon, M., Wu, Y., Jeang, K.T., and Wattel, E. (2003). Inactivation of hTERT transcription by Tax. *Oncogene* 22, 3734-3741.

Gachon, F., Thebault, S., Peleraux, A., Devaux, C., and Mesnard, J.M. (2000). Molecular interactions involved in the transactivation of the human T-cell leukemia virus type 1 promoter mediated by Tax and CREB-2 (ATF-4). *Mol Cell Biol* 20, 3470-3481.

Gallo, R.C. (2005). History of the discoveries of the first human retroviruses: HTLV-1 and HTLV-2. *Oncogene* 24, 5926-5930.

Gaudray, G., Gachon, F., Basbous, J., Biard-Piechaczyk, M., Devaux, C., and Mesnard, J.M. (2002). The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription. *J Virol* 76, 12813-12822.

Geleziunas, R., Ferrell, S., Lin, X., Mu, Y., Cunningham, E.T., Jr., Grant, M., Connelly, M.A., Hambor, J.E., Marcu, K.B., and Greene, W.C. (1998). Human T-cell leukemia virus type 1 Tax induction of NF-kappaB involves activation of the IkappaB kinase alpha (IKKalpha) and IKKbeta cellular kinases. *Mol Cell Biol* 18, 5157-5165.

Gilbert, S.F. (2000). *Developmental Biology*. Sunderland (MA), Sinauer Associates, Inc.

Halin, M., Douceron, E., Clerc, I., Journo, C., Ko, N.G., Landry, S., Murphy, E., Mesnard, J.M., Barbeau, B., and Mahieux, R. (2008). Human T-cell Leukemia Virus Type 2 produces a spliced antisense transcript encoding a protein that lacks a bZIP domain but inhibits Tax2-mediated transcription. *Blood Submitted*.

Hivin, P., Arpin-Andre, C., Clerc, I., Barbeau, B., and Mesnard, J.M. (2006). A modified version of a Fos-associated cluster in HBZ affects Jun transcriptional potency. *Nucleic Acids Res* 34, 2761-2772.

Hivin, P., Basbous, J., Raymond, F., Henaff, D., Arpin-Andre, C., Robert-Hebmann, V., Barbeau, B., and Mesnard, J.M. (2007). The HBZ-SP1 isoform of human T-cell leukemia virus type I represses JunB activity by sequestration into nuclear bodies. *Retrovirology* 4, 14.

Hivin, P., Frederic, M., Arpin-Andre, C., Basbous, J., Gay, B., Thebault, S., and Mesnard, J.M. (2005). Nuclear localization of HTLV-I bZIP factor (HBZ) is mediated by three distinct motifs. *J Cell Sci* 118, 1355-1362.

Igakura, T., Stinchcombe, J.C., Goon, P.K., Taylor, G.P., Weber, J.N., Griffiths, G.M., Tanaka, Y., Osame, M., and Bangham, C.R. (2003). Spread of HTLV-I between lymphocytes by virus-induced polarization of the cytoskeleton. *Science* 299, 1713-1716.

Jin, D.Y., Spencer, F., and Jeang, K.T. (1998). Human T cell leukemia virus type 1 oncoprotein Tax targets the human mitotic checkpoint protein MAD1. *Cell* 93, 81-91.

Jones, K.S., Petrow-Sadowski, C., Huang, Y.K., Bertolette, D.C., and Ruscetti, F.W. (2008). Cell-free HTLV-1 infects dendritic cells leading to transmission and transformation of CD4(+) T cells. *Nat Med* 14, 429-436.

Kalyanaraman, V.S., Sarngadharan, M.G., Robert-Guroff, M., Miyoshi, I., Golde, D., and Gallo, R.C. (1982). A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia. *Science* 218, 571-573.

Kashanchi, F., and Brady, J.N. (2005). Transcriptional and post-transcriptional gene regulation of HTLV-1. *Oncogene* 24, 5938-5951.

Kinet, S., Swainson, L., Lavanya, M., Mongellaz, C., Montel-Hagen, A., Craveiro, M., Manel, N., Battini, J.L., Sitbon, M., and Taylor, N. (2007). Isolated receptor binding domains of HTLV-1 and HTLV-2 envelopes bind Glut-1 on activated CD4+ and CD8+ T cells. *Retrovirology* 4, 31.

Kuhlmann, A.S., Villaudy, J., Gazzolo, L., Castellazzi, M., Mesnard, J.M., and Duc Dodon, M. (2007). HTLV-1 HBZ cooperates with JunD to enhance transcription of the human telomerase reverse transcriptase gene (hTERT). *Retrovirology* 4, 92.

Lamsoul, I., Lodewick, J., Lebrun, S., Brasseur, R., Burny, A., Gaynor, R.B., and Bex, F. (2005). Exclusive ubiquitination and sumoylation on overlapping lysine residues mediate NF-kappaB activation by the human T-cell leukemia virus tax oncoprotein. *Mol Cell Biol* 25, 10391-10406.

Landais, S., Landry, S., Legault, P., and Rassart, E. (2007). Oncogenic potential of the miR-106-363 cluster and its implication in human T-cell leukemia. *Cancer Res* 67, 5699-5707.

Landry, S., Halin, M., Lefort, S., Audet, B., Vaquero, C., Mesnard, J.M., and Barbeau, B. (2007). Detection, characterization and regulation of antisense transcripts in HIV-1. *Retrovirology* 4, 71.

Landry, S., Halin, M., Vargas, A., Lemasson, I., Mesnard, J.M., and Barbeau, B. (2008). Upregulation of HTLV-1 antisense transcription by the viral Tax protein. *Journal of Virology* *submitted*.

Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240, 1759-1764.

Larocca, D., Chao, L.A., Seto, M.H., and Brunck, T.K. (1989). Human T-cell leukemia virus minus strand transcription in infected T-cells. *Biochem Biophys Res Commun* 163, 1006-1013.

Lee, W., and Keller, E.B. (1991). Regulatory elements mediating transcription of the human Ha-ras gene. *J Mol Biol* 220, 599-611.

Lemasson, I., Lewis, M.R., Polakowski, N., Hivin, P., Cavanagh, M.H., Thebault, S., Barbeau, B., Nyborg, J.K., and Mesnard, J.M. (2007). Human T-cell leukemia virus type 1 (HTLV-1) bZIP protein interacts with the cellular transcription factor CREB to inhibit HTLV-1 transcription. *J Virol* 81, 1543-1553.

Lemoine, F.J., Kao, S.Y., and Marriott, S.J. (2000). Suppression of DNA repair by HTLV type 1 Tax correlates with Tax trans-activation of proliferating cell nuclear antigen gene expression. *AIDS Res Hum Retroviruses* 16, 1623-1627.

- Li, M., and Green, P.L. (2007). Detection and quantitation of HTLV-1 and HTLV-2 mRNA species by real-time RT-PCR. *J Virol Methods* 142, 159-168.
- Mahieux, R., and Gessain, A. (2005). [New human retroviruses: HTLV-3 and HTLV-4]. *Med Trop (Mars)* 65, 525-528.
- Mahieux, R., Pise-Masison, C.A., Lambert, P.F., Nicot, C., De Marchis, L., Gessain, A., Green, P., Hall, W., and Brady, J.N. (2000). Differences in the ability of human T-cell lymphotropic virus type 1 (HTLV-1) and HTLV-2 tax to inhibit p53 function. *J Virol* 74, 6866-6874.
- Manel, N., Kim, F.J., Kinet, S., Taylor, N., Sitbon, M., and Battini, J.L. (2003). The ubiquitous glucose transporter GLUT-1 is a receptor for HTLV. *Cell* 115, 449-459.
- Marriott, S.J., and Semmes, O.J. (2005). Impact of HTLV-I Tax on cell cycle progression and the cellular DNA damage repair response. *Oncogene* 24, 5986-5995.
- Matsumoto, J., Ohshima, T., Isono, O., and Shimotohno, K. (2005). HTLV-1 HBZ suppresses AP-1 activity by impairing both the DNA-binding ability and the stability of c-Jun protein. *Oncogene* 24, 1001-1010.
- Matsumoto, K., Shibata, H., Fujisawa, J.I., Inoue, H., Hakura, A., Tsukahara, T., and Fujii, M. (1997). Human T-cell leukemia virus type 1 Tax protein transforms rat fibroblasts via two distinct pathways. *J Virol* 71, 4445-4451.
- Matsuoka, M., and Jeang, K.T. (2007). Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. *Nat Rev Cancer* 7, 270-280.
- Meertens, L., Chevalier, S., Weil, R., Gessain, A., and Mahieux, R. (2004a). A 10-amino acid domain within human T-cell leukemia virus type 1 and type 2 tax protein sequences is responsible for their divergent subcellular distribution. *J Biol Chem* 279, 43307-43320.
- Meertens, L., Pise-Masison, C., Quere, N., Brady, J., Gessain, A., and Mahieux, R. (2004b). Utilization of the CBP but not the p300 co-activator by human T-lymphotropic virus type-2 Tax for p53 inhibition. *Oncogene* 23, 5447-5458.
- Mesnard, J.M., Barbeau, B., and Devaux, C. (2006). HBZ, a new important player in the mystery of adult T-cell leukemia. *Blood* 108, 3979-3982.
- Michael, N.L., Vahey, M.T., d'Arcy, L., Ehrenberg, P.K., Mosca, J.D., Rappaport, J., and Redfield, R.R. (1994). Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat. *J Virol* 68, 979-987.

Miller, R.H. (1988). Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand. *Science* 239, 1420-1422.

Mitchell, R.S., Beitzel, B.F., Schroder, A.R., Shinn, P., Chen, H., Berry, C.C., Ecker, JR., and Bushman, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PloS Biol.* 2, 1127-1137.

Miyake, H., Suzuki, T., Hirai, H., and Yoshida, M. (1999). Trans-activator Tax of human T-cell leukemia virus type 1 enhances mutation frequency of the cellular genome. *Virology* 253, 155-161.

Miyazaki, M., Yasunaga, J., Taniguchi, Y., Tamiya, S., Nakahata, T., and Matsuoka, M. (2007). Preferential selection of human T-cell leukemia virus type 1 provirus lacking the 5' long terminal repeat during oncogenesis. *J Virol* 81, 5714-5723.

Miyoshi, I., Kubonishi, I., Yoshimoto, S., Akagi, T., Ohtsuki, Y., Shiraishi, Y., Nagata, K., and Hinuma, Y. (1981). Type C virus particles in a cord T-cell line derived by co-cultivating normal human cord leukocytes and human leukaemic T cells. *Nature* 294, 770-771.

Murata, K., Hayashibara, T., Sugahara, K., Uemura, A., Yamaguchi, T., Harasawa, H., Hasegawa, H., Tsuruda, K., Okazaki, T., Koji, T., *et al.* (2006). A novel alternative splicing isoform of human T-cell leukemia virus type 1 bZIP factor (HBZ-SI) targets distinct subnuclear localization. *J Virol* 80, 2495-2505.

Peeters, A., Lambert, P.F., and Deacon, N.J. (1996). A fourth Spl site in the human immunodeficiency virus type 1 long terminal repeat is essential for negative-sense transcription. *J Virol* 70, 6665-6672.

Poiesz, B.J., Poiesz, M.J., and Choi, D. (2003). The human T-cell lymphoma/leukemia viruses. *Cancer Invest* 21, 253-277.

Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D., and Gallo, R.C. (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* 77, 7415-7419.

Poiesz, B.J., Ruscetti, F.W., Reitz, M.S., Kalyanaraman, V.S., and Gallo, R.C. (1981). Isolation of a new type C retrovirus (HTLV) in primary uncultured cells of a patient with Sezary T-cell leukaemia. *Nature* 294, 268-271.

Ross, T.M., Minella, A.C., Fang, Z.Y., Pettiford, S.M., and Green, P.L. (1997). Mutational analysis of human T-cell leukemia virus type 2 Tax. *J Virol* 71, 8912-8917.

Roucoux, D.F., and Murphy, E.L. (2004). The epidemiology and disease outcomes of human T-lymphotropic virus type II. *AIDS Rev* 6, 144-154.

Sakamoto, K.M., Nimer, S.D., Rosenblatt, J.D., and Gasson, J.C. (1992). HTLV-I and HTLV-II tax trans-activate the human EGR-1 promoter through different cis-acting sequences. *Oncogene* 7, 2125-2130.

Satou, Y., Yasunaga, J., Yoshida, M., and Matsuoka, M. (2006). HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. *Proc Natl Acad Sci U S A* 103, 720-725.

Sawadogo, M., and Roeder, R.G. (1984). Energy requirement for specific transcription initiation by the human RNA polymerase II system. *J Biol Chem* 259, 5321-5326.

Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R. and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110, 521-529.

Semmes, O.J., Majone, F., Cantemir, C., Turchetto, L., Hjelle, B., and Jeang, K.T. (1996). HTLV-I and HTLV-II Tax: differences in induction of micronuclei in cells and transcriptional activation of viral LTRs. *Virology* 217, 373-379.

Shimotohno, K., Takahashi, Y., Shimizu, N., Gojobori, T., Golde, D.W., Chen, I.S., Miwa, M., and Sugimura, T. (1985). Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: an open reading frame for the protease gene. *Proc Natl Acad Sci U S A* 82, 3101-3105.

Shimoyama, M. (1991). Diagnostic criteria and classification of clinical subtypes of adult T-cell leukaemia-lymphoma. A report from the Lymphoma Study Group (1984-87). *Br J Haematol* 79, 428-437.

Silverman, L.R., Phipps, A.J., Montgomery, A., Ratner, L., and Lairmore, M.D. (2004). Human T-cell lymphotropic virus type 1 open reading frame II-encoded p30II is required for in vivo replication: evidence of in vivo reversion. *J Virol* 78, 3837-3845.

Starr, D.B., and Hawley, D.K. (1991). TFIID binds in the minor groove of the TATA box. *Cell* 67, 1231-1240.

Switzer, W.M., Qari, S.H., Wolfe, N.D., Burke, D.S., Folks, T.M., and Heneine, W. (2006). Ancient origin and molecular features of the novel human T-lymphotropic virus type 3 revealed by complete genome analysis. *J Virol* 80, 7427-7438.

Tagieva, N.E., and Vaquero, C. (1997). Expression of naturally occurring antisense RNA inhibits human immunodeficiency virus type 1 heterologous strain replication. *J Gen Virol* 78 (Pt 10), 2503-2511.

Tanaka, K., Shiraishi, K., Sakamoto, A., Okubo, Y., Fukuda, Y., Fukahori, S., Jojima, H., Osabe, S., Honda, J., Natori, H., *et al.* (1996). [Clinical significance of soluble interleukin-2 receptor levels in the cerebrospinal fluid in patients with adult T-cell leukemia complicated with meningeal infiltration]. *Gan To Kagaku Ryoho* 23, 863-869.

Thebault, S., Basbous, J., Hivin, P., Devaux, C., and Mesnard, J.M. (2004). HBZ interacts with JunD and stimulates its transcriptional activity. *FEBS Lett* 562, 165-170.

Tripp, A., Liu, Y., Sieburg, M., Montalbano, J., Wrzesinski, S., and Feuer, G. (2003). Human T-cell leukemia virus type 1 tax oncoprotein suppression of multilineage hematopoiesis of CD34+ cells in vitro. *J Virol* 77, 12152-12164.

Uchida, N., Otsuka, T., Arima, F., Shigematsu, H., Fukuyama, T., Maeda, M., Sugio, Y., Itoh, Y., and Niho, Y. (1999). Correlation of telomerase activity with development and progression of adult T-cell leukemia. *Leuk Res* 23, 311-316.

Uren, A.G., Kool, J., Berns, A., and Van Lohuizen, M. (2005). Retroviral insertional mutagenesis: past, present and future. *Oncogene* 24, 7656-7672.

Usui, T., Yanagihara, K., Tsukasaki, K., Murata, K., Hasegawa, H., Yamada, Y., and Kamihira, S. (2008). Characteristic expression of HTLV-1 basic zipper factor (HBZ) transcripts in HTLV-1 provirus-positive cells. *Retrovirology* 5, 34.

Van Dooren, S., Salemi, M., and Vandamme, A.M. (2001). Dating the origin of the African human T-cell lymphotropic virus type-i (HTLV-I) subtypes. *Mol Biol Evol* 18, 661-671.

Van, P.L., Yim, K.W., Jin, D.Y., Dapolito, G., Kurimasa, A., and Jeang, K.T. (2001). Genetic evidence of a role for ATM in functional interaction between human T-cell leukemia virus type 1 Tax and p53. *J Virol* 75, 396-407.

Vanhee-Brossollet, C., Thoreau, H., Serpente, N., D'Auriol, L., Levy, J.P., and Vaquero, C. (1995). A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals. *Virology* 206, 196-202.

Vanhee-Brossollet, C., and Vaquero, C. (1998). Do natural antisense transcripts make sense in eukaryotes? *Gene* 211, 1-9.

Verdonck, K., Gonzalez, E., Henostroza, G., Nabeta, P., Llanos, F., Cornejo, H., Vanham, G., Seas, C., and Gotuzzo, E. (2007). HTLV-1 infection is frequent among outpatients with pulmonary tuberculosis in northern Lima, Peru. *Int J Tuberc Lung Dis* 11, 1066-1072.

Wolfe, N.D., Heneine, W., Carr, J.K., Garcia, A.D., Shanmugam, V., Tamoufe, U., Torimiro, J.N., Prosser, A.T., Lebreton, M., Mpoudi-Ngole, E., *et al.* (2005). Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. *Proc Natl Acad Sci U S A* 102, 7994-7999.

Yoshida, M., Miyoshi, I., and Hinuma, Y. (1982). Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc Natl Acad Sci U S A* 79, 2031-2035.

Yoshida, M., Satou, Y., Yasunaga, J.I., Fujisawa, J.I., and Matsuoka, M. (2008). Transcriptional control of spliced and unspliced HTLV-1 bZIP factor gene. *J Virol*.

Younis, I., Yamamoto, B., Phipps, A., and Green, P.L. (2005). Human T-cell leukemia virus type 1 expressing nonoverlapping tax and rex genes replicates and immortalizes primary human T lymphocytes but fails to replicate and persist in vivo. *J Virol* 79, 14473-14481.